

## SCORE Search Results for Application 09821877

[Score Home  
Page](#)[Retrieve Application  
List](#)[SCORE System  
Overview](#)[SCORE  
FAQ](#)[Comments /  
Suggestions](#)

This page gives you a list of all the Search Results. Use this page to obtain specific Search Result information.

[View version list for this application](#)

[Go Back to previous page](#)

### Item Listing Version# 1

Item Name	Download Content
<a href="#">us-09-821-877-1.rge</a>	<a href="#">Download</a>
<a href="#">us-09-821-877-1.rng</a>	<a href="#">Download</a>
<a href="#">us-09-821-877-1.rni</a>	<a href="#">Download</a>
<a href="#">us-09-821-877-1.rnpbm</a>	<a href="#">Download</a>
<a href="#">us-09-821-877-1.rnpbn</a>	<a href="#">Download</a>
<a href="#">us-09-821-877-1.rnpm</a>	<a href="#">Download</a>
<a href="#">us-09-821-877-1.rnpn</a>	<a href="#">Download</a>
<a href="#">us-09-821-877-1.rst</a>	<a href="#">Download</a>

**SCORE Search Results Details for Application 09821877 and Search Result us-09-821-877-1.rge.**

[Score Home](#)  
[Page](#)

## Retrieve Application List

## SCORE System Overview

SCORE  
FAQ

Comments /  
Suggestions

This page gives you Search Results detail for the Application 09821877 and Search Result us-09-821-877-1.rge.

start

[Go Back to previous page](#)

GenCore version 5.1.9  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

```
Run on:      July 22, 2006, 03:40:49 ; Search time 10969 Seconds
              (without alignments)
              6885.031 Million cell updates/sec
```

```
Title:          US-09-821-877-1
Perfect score:  1181
Sequence:       1 atgggggcagaatctttccac.....tacattttaaacctaataaa 1181
```

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 6366136 seqs, 31973710525 residues

Total number of hits satisfying chosen parameters: 12732272

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

```
Post-processing: Minimum Match 0%
                  Maximum Match 100%
                  Listing first 45 summaries
```

```
Database : GenEmbl:*
1: gb_env:*
2: gb_pat:*
3: gb_ph:*
4: gb_pl:*
5: gb_pr:*
6: gb_ro:*
7: gb_sts:*
8: gb_sy:*
9: gb_un:*
10: gb_vi:*
11: gb_ov:*
12: gb_htg:*
13: gb_in:*
14: gb_om:*
15: gb_ba:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1136.2	96.2	1186	10	AY230122	AY230122 Hepatitis
2	1133.2	96.0	1170	10	AB104717	AB104717 Hepatitis
3	1133	95.9	2612	2	E00007	E00007 DNA coding
4	1133	95.9	2743	2	A08967	A08967 Hepatitis B
5	1133	95.9	2743	10	HPBADYW	J02202 human hepat
6	1131.6	95.8	1170	10	AB104723	AB104723 Hepatitis
7	1128.4	95.5	1170	10	AB104715	AB104715 Hepatitis
8	1128.2	95.5	8007	2	BD181818	BD181818 Hepatitis
9	1126.8	95.4	1170	10	AY603461	AY603461 Hepatitis
10	1126.6	95.4	3207	8	DQ219811	DQ219811 Synthetic
11	1125.2	95.3	1170	10	AB104722	AB104722 Hepatitis
12	1125.2	95.3	1170	10	AY576427	AY576427 Hepatitis
13	1125.2	95.3	1170	10	AY603460	AY603460 Hepatitis
14	1125.2	95.3	1170	10	AY603464	AY603464 Hepatitis
15	1125	95.3	1186	10	AY230128	AY230128 Hepatitis
16	1123	95.1	1171	10	DQ131119	DQ131119 Hepatitis
17	1122	95.0	1170	10	AY603456	AY603456 Hepatitis
18	1122	95.0	1170	10	AY603459	AY603459 Hepatitis
19	1121.8	95.0	1193	10	AY230125	AY230125 Hepatitis
20	1121.8	95.0	3187	10	AY236162	AY236162 Hepatitis
21	1121.8	95.0	3187	10	DQ336678	DQ336678 Hepatitis
22	1120.4	94.9	1170	10	AY603451	AY603451 Hepatitis
23	1120.4	94.9	1170	10	AY603457	AY603457 Hepatitis
24	1120.2	94.9	3187	10	DQ304548	DQ304548 Hepatitis
25	1120.2	94.9	3187	10	DQ336675	DQ336675 Hepatitis
26	1120.2	94.9	3187	10	DQ336679	DQ336679 Hepatitis
27	1118.8	94.7	1170	10	AB104719	AB104719 Hepatitis
28	1118.8	94.7	1170	10	AY603450	AY603450 Hepatitis
29	1118.8	94.7	1170	10	AY603454	AY603454 Hepatitis
30	1118.8	94.7	1170	10	AY603462	AY603462 Hepatitis
31	1118.6	94.7	3187	10	DQ304549	DQ304549 Hepatitis
32	1118.6	94.7	3187	10	DQ304551	DQ304551 Hepatitis
33	1118.6	94.7	3187	10	DQ336674	DQ336674 Hepatitis
34	1117.2	94.6	1170	10	AY603465	AY603465 Hepatitis
35	1117.2	94.6	1201	2	AR011346	AR011346 Sequence
36	1117.2	94.6	1201	2	I17984	I17984 Sequence 21
37	1117.2	94.6	1285	2	AR011345	AR011345 Sequence
38	1117.2	94.6	1285	2	I17983	I17983 Sequence 21
39	1117.2	94.6	2342	2	A32618	A32618 Synthetic c
40	1117	94.6	2637	10	AY902771	AY902771 Hepatitis
41	1117	94.6	3187	10	DQ304547	DQ304547 Hepatitis
42	1117	94.6	3187	10	DQ304550	DQ304550 Hepatitis
43	1117	94.6	3187	10	DQ336676	DQ336676 Hepatitis
44	1115.6	94.5	1170	10	AB074844	AB074844 Hepatitis
45	1115.4	94.4	3187	10	DQ336677	DQ336677 Hepatitis

## ALIGNMENTS

## RESULT 1

AY230122

LOCUS AY230122 1186 bp DNA linear VRL 23-NOV-2003

DEFINITION Hepatitis B virus isolate case 18 non-tumor large surface protein, middle surface protein, and small surface protein genes, complete cds.

ACCESSION AY230122

VERSION AY230122.1 GI:38374274

KEYWORDS .

SOURCE Hepatitis B virus

ORGANISM Hepatitis B virus  
 Viruses; Retro-transcribing viruses; Hepadnaviridae;  
 Orthohepadnavirus.

REFERENCE 1 (bases 1 to 1186)  
 AUTHORS Raimondo,G., Pollicino,T. and Raffa,G.  
 TITLE Occult HBV in liver cancer  
 JOURNAL Unpublished

REFERENCE 2 (bases 1 to 1186)  
 AUTHORS Raimondo,G., Pollicino,T. and Raffa,G.  
 TITLE Direct Submission  
 JOURNAL Submitted (04-FEB-2003) Internal Medicine, University of Messina,  
 via Consolare Valeria, Messina 98124, Italy

FEATURES Location/Qualifiers

source 1. .1186  
 /organism="Hepatitis B virus"  
 /virion  
 /mol\_type="genomic DNA"  
 /isolate="case 18 non-tumor"  
 /db\_xref="taxon:10407"

misc\_feature 1. .1170  
 /note="similar to large surface protein"

misc\_feature 325. .1170  
 /note="similar to middle surface protein"

misc\_feature 490. .1170  
 /note="similar to small surface protein"

## ORIGIN

Query Match 96.2%; Score 1136.2; DB 10; Length 1186;  
 Best Local Similarity 97.6%; Pred. No. 0;  
 Matches 1153; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

Qy	1	ATGGGGCAGAATCTTTCCACCAGCAATCCTCTGGGATTCTTTCCCGACCACCAGTTGGAT	60
Db	1	ATGGGGCAGAATCTTTCCACCAGCAATCCTCTGGGATTCTTTCCCGACCACCAGTTGGAT	60
Qy	61	CCAGCCTTCAGAGCAAACACCAACAATCCAGATTGGGACTTCAATCCCAACAAGGACACC	120
Db	61	CCAGCCTTCAGAGCAAACACCGCAAATCCAGATTGGGACTTCAATCCCAACAAGGACACC	120
Qy	121	TGGCCAGACGCCAACAAAGGTAGGAGCTGGAGCATTCGGACTGGGGTTCACCCACCGCAC	180
Db	121	TGGCCAGACGCCAACAAAGGTAGGAGCTGGAGCATTCGGGCTGGGATTACCCACCGCAC	180
Qy	181	GGAGGCCTTTTGGGGTGGAGCCCTCAGGCTCAGGGCATAACACAAACCTTGCCAGCAAAT	240
Db	181	GGAGGCCTTTTGGGGTGGAGCCCTCAGGCTCAGGGCATACTACAAACCTTGCCAGCAAAT	240
Qy	241	CCGCCTCCTGCTTCCACCAATCGCCAGTCAGGAAGGCAGCCTACCCCGCTGTCTCCACCT	300
Db	241	CCGCCTCCTGCTTCTACCAATCGCCAGTCAGGAAGGCAGCCTACCCCTCTGTCTCCACCT	300
Qy	301	TTGAGAAACACTCATCCTCAAGCCATGCAGTGGAAGTCCACAACCTTCCACCAAACCTCTG	360
Db	301	TTGAGAAACACTCATCCTCAGGCCATGCAGTGGAAGTCCACAACCTTCCACCAAACCTCTG	360
Qy	361	CAAGATCCCAGAGTGAGAGGTCTGTATTTCCCTGCTGGTGGCTCCAGTTCAGGAACAGTA	420
Db	361	CAAGATCCCAGAGTGAGAGGCCTGTATTTCCCTGCTGGTGGCTCCAGTTCAGGAACAGTA	420
Qy	421	AACCCTGTTCCGACTACTGTCTCTCCCATATCGTCAATCTTCTCGAGGATTGGGGACCCT	480
Db	421	AACCCTGTTCCGACTACTGTCTCTCCCATATCGTCAATCTTCTCGAGGATTGGGGACCCT	480
Qy	481	GCGCGGAACATGGAGAACATCACATCAGGATTCTAGGACCCCTGCTCGTGTACAGGCG	540

http://es/ScoreAccessWeb/GetItem.action?AppId=09821877&seqId=566763&ItemName=us... 8/4/06

delta viruses in Egypt  
 JOURNAL J. Med. Virol. 70 (4), 529-536 (2003)  
 PUBMED 12794714  
 REFERENCE 2 (bases 1 to 1170)  
 AUTHORS Suzuki,S., Saady,N., Sugauchi,F., Orito,E., Agha,S. and Mizokami,M.  
 TITLE Direct Submission  
 JOURNAL Submitted (26-FEB-2003) Seiji Suzuki, Nagoya City University  
 Graduate School, Department of clinical Molecular Informative  
 Medicine; Mizuho, Nagoya, Aichi 467-8601, Japan  
 (E-mail:seijis@med.nagoya-cu.ac.jp, Tel:81-52-853-8292,  
 Fax:81-52-842-0021)

FEATURES Location/Qualifiers  
 source 1. .1170  
 /organism="Hepatitis B virus"  
 /mol\_type="genomic DNA"  
 /isolate="EG80"  
 /db\_xref="taxon:10407"  
 /country="Egypt"  
 /note="genotype D"  
 gene 1. .1170  
 /gene="s"  
 CDS 1. .1170  
 /gene="s"  
 /codon\_start=1  
 /product="pre-S and S protein"  
 /protein\_id="BAC78512.1"  
 /db\_xref="GI:32261195"  
 /translation="MGQNLSTSNPLGFFPDHQLDPAFRANTANPDWDFNPNKDTWPDAN  
 NKVGAGAGFLGFTPPHGGLLGWSPPQAQGIQLTPANPPPASTNRQSGRQPTPLSPPLR  
 NTHPQAMQWNSTTFHQTLQDPRVRGLYFPAGSSSGTVNPVPTTVSPISSIFSRIGDP  
 ALNMENITSGFLGPLLVLQAGFLLTRILTIPQSLDSWWTSLNFLGGTTVCLGQNSQS  
 PTSNHSPTSCPPTCPGYRWMCLRRFIIFLFILLCLIFLLVLLDYQGMLPVCPLIPGS  
 STTSTGPCRTCTTPAQGTSMPSCCCTKPLDGNCTCIPIPSSWAFGKFLWEWASARFS  
 WLSLLVPFVQWVGLSPTVWLSVIWMMWYWGPSLYSILSPFLPLLPIFFCLWVYI"

## ORIGIN

Query Match 96.0%; Score 1133.2; DB 10; Length 1170;  
 Best Local Similarity 98.0%; Pred. No. 0;  
 Matches 1147; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

Qy	1	ATGGGGCAGAATCTTTCCACCAGCAATCCTCTGGGATTCTTTCCCGACCACCAGTTGGAT	60
Db	1	ATGGGGCAGAATCTTTCCACCAGCAATCCTCTGGGATTCTTTCCCGACCACCAGTTGGAT	60
Qy	61	CCAGCCTTCAGAGCAAACACCAACAATCCAGATTGGGACTTCAATCCCAACAAGGACACC	120
Db	61	CCAGCCTTCAGAGCAAACACCGCAAATCCAGATTGGGACTTCAATCCCAACAAGGACACC	120
Qy	121	TGGCCAGACGCCAACAAGGTAGGAGCTGGAGCATTCGGGACTGGGGTTACCCCCACCGCAC	180
Db	121	TGGCCAGACGCCAACAAGGTAGGAGCTGGAGCATTCGGGCTGGGATTACCCCCACCGCAC	180
Qy	181	GGAGGCCTTTTGGGGTGGAGCCCTCAGGCTCAGGGCATAACACAAACCTTGCCAGCAAAT	240
Db	181	GGAGGCCTTTTGGGGTGGAGCCCTCAGGCTCAGGGCATACTACAAACCTTGCCAGCAAAT	240
Qy	241	CCGCCTCTGCTTCCACCAATCGCCAGTCAGGAAGGCAGCCTACCCCGCTGTCTCCACCT	300
Db	241	CCGCCTCTGCTTCTACCAATCGCCAGTCAGGAAGGCAGCCTACCCCTCTGTCTCCACCT	300
Qy	301	TTGAGAAACACTCATCCTCAAGCCATGCAGTGGAAGTCCACAACCTTTCCACCAAACCTCTG	360
Db	301	CTGAGAAACACTCATCCTCAGGCCATGCAGTGGAAGTCCACAACCTTCCACCAAACCTCTG	360

```
Qy      361 CAAGATCCCAGAGTGAGAGGTCTGTATTTCCCTGCTGGTGGCTCCAGTTCAGGAACAGTA 420
      |||
Db      361 CAAGATCCCAGAGTGAGAGGCCCTGTATTTCCCTGCTGGTGGCTCCAGTTCAGGAACAGTA 420

Qy      421 AACCTGTTCCGACTACTGTCTCTCCCATATCGTCAATCTTCTCGAGGATTGGGGACCCT 480
      |||
Db      421 AACCTGTTCCGACTACTGTCTCTCCCATATCGTCAATCTTCTCGAGGATTGGGGACCCT 480

Qy      481 GCGCGGAACATGGAGAACATCACATCAGGATTCTAGGACCCCTGCTCGTGTACAGGCG 540
      |||
Db      481 GCGTTGAACATGGAGAACATCACATCAGGATTCTAGGACCCCTGCTCGTGTACAGGCG 540

Qy      541 GGGTTTTTCTTGTTGACAAGAATCCTCACAATACCGCAGAGTCTAGACTCGTGGTGGACT 600
      |||
Db      541 GGGTTTTTCTTGTTGACAAGAATCCTCACAATACCGCAGAGTCTAGACTCGTGGTGGACT 600

Qy      601 TCTCTCAATTTTCTAGGGGGAACCTACCGTGTGTCTTGGCCAAAATTTCGCAGTCCCCAACC 660
      |||
Db      601 TCTCTCAATTTTCTAGGGGGAACCTACCGTGTGTCTTGGCCAAAATTTCGCAGTCCCCAACC 660

Qy      661 TCCAATCACTACCAACCTCCTGTCTCCAACCTTGTCTGGTTATCGCTGGATGTGTCTG 720
      |||
Db      661 TCCAATCACTACCAACCTCCTGTCTCCAACCTTGTCTGGTTATCGCTGGATGTGTCTG 720

Qy      721 CGGCGTTTTATCATCTTCTCTTCATCCTGCTGCTATGCCTCATCTTCTTGTTGGTTCTT 780
      |||
Db      721 CGGCGTTTTATCATCTTCTCTTCATCCTGCTGCTATGCCTCATCTTCTTGTTGGTTCTT 780

Qy      781 CTGGACTATCAAGGTATGTTGCCGTTTGTCTCTAATTCCAGGATCTTCAACCACCAGC 840
      |||
Db      781 CTGGACTATCAAGGTATGTTGCCGTTTGTCTCTAATTCCAGGATCTTCAACCACCAGC 840

Qy      841 ACGGGACCATGCAGAGCCTGCACGACTCCTGCTCAAGGAACCTCTATGTATCCCTCCTGT 900
      |||
Db      841 ACGGGACCATGCAGAACCTGCACGACTCCTGCTCAAGGAACCTCTATGTATCCCTCCTGT 900

Qy      901 TGCTGTACAAAACCTTCGGATGGAACTGCACCTGTATTCCCATCCCATCATCCTGGGCT 960
      |||
Db      901 TGCTGTACAAAACCTTTGGACGGAAATGCACCTGTATTCCCATCCCATCATCCTGGGCT 960

Qy      961 TTCGGAATAATCCTATGGGAGTGGGCCTCAGCCCGTTTCTCCTGGCTCAGTTTACTAGTG 1020
      |||
Db      961 TTCGGAATAATCCTATGGGAGTGGGCCTCAGCCCGTTTCTCCTGGCTCAGTTTACTAGTG 1020

Qy      1021 CCATTTGTTTCAGTGGTTTCGTAGGGCTTCCCCCACTGTTTGGCTTTCAGTTATATGGATG 1080
      |||
Db      1021 CCATTTGTTTCAGTGGTTTCGTAGGGCTTCCCCCACTGTTTGGCTTTCAGTTATATGGATG 1080

Qy      1081 ATGTTGTACTGGGGGCCAAGTCTGTACACCATCTTGAGTCCCTTTTTACCGCTGTTACCA 1140
      |||
Db      1081 ATGTGGTATTGGGGGCCAAGTCTGTACAGCATCTTGAGTCCCTTTTTACCGCTGTTACCA 1140

Qy      1141 ATTTTCTTTTGTCTTTGGGTATACATTTAA 1170
      |||
Db      1141 ATTTTCTTTTGTCTTTGGGTATACATTTAA 1170
```

## RESULT 3

E00007

LOCUS E00007 2612 bp DNA linear PAT 29-SEP-1997

DEFINITION DNA coding of HBV antigen.

ACCESSION E00007

VERSION E00007.1 GI:2168318

KEYWORDS JP 1980104887-A/1.

SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominidae; Homo.  
REFERENCE 1 (bases 1 to 2612)  
AUTHORS Kenesu,M. and Haintsu,E.S.  
TITLE REARRANGED DNA MOLECULE AND METHOD  
JOURNAL Patent: JP 1980104887-A 1 11-AUG-1980;  
BIOGEN NV  
COMMENT OS human  
PN JP 1980104887-A/1  
PD 11-AUG-1980  
PF 20-DEC-1979 JP 1979164945  
PR 22-DEC-1978 GB 78 49907, 27-DEC-1978 GB 78 50039, PR  
01-NOV-1979 GB 79 7937910  
PI KENESU MAREE, HAINTSU ERUNSUTO SHIYARAA  
PC C12N15/00,C07H21/04,C12N1/00,C12P19/34,C12P21/02,C12Q1/00, PC  
C12Q1/70//  
PC C12R1/125,C12R1/19,C12R1/38,C12R1/645;  
CC strandedness: Double;  
CC topology: Linear;  
CC hypothetical: No;  
CC anti-sense: No;  
CC \*source: tissue\_type=leukocyte;  
FH Key Location/Qualifiers  
FH  
FT sig\_peptide 1. .87  
FT mat\_peptide 88. .636  
FT /product='HBV nucleus antigen' FT  
mat\_peptide 1524. .2201  
FT /product='HBV surface antigen' FT CDS  
1. .639  
FT /product='HBV nucleus antigen' FT CDS  
1524. .2204  
FT /product='HBV surface antigen'.  
FEATURES Location/Qualifiers  
source 1. .2612  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"

## ORIGIN

Query Match 95.9%; Score 1133; DB 2; Length 2612;  
Best Local Similarity 97.5%; Pred. No. 0;  
Matches 1151; Conservative 0; Mismatches 30; Indels 0; Gaps 0;

```
Qy      1 ATGGGGCAGAATCTTTCCACCAGCAATCCTCTGGGATTCTTTCCCGACCACCAGTTGGAT 60
          |||
Db      1035 ATGGGGCAGAATCTTTCCACCAGCAATCCTCTGGGATTCTTTCCCGACCACCAGTTGGAT 1094

Qy      61 CCAGCCTTCAGAGCAAACACCAACAATCCAGATTGGGACTTCAATCCCAACAAGGACACC 120
          |||
Db      1095 CCAGCCTTCAGAGCAAACACCAACAATCCAGATTGGGACTTCAATCCCAACAAGGACACC 1154

Qy      121 TGGCCAGACGCCAACAAGGTAGGAGCTGGAGCATTCGGACTGGGGTTACCCCCACCGCAC 180
          |||
Db      1155 TGGCCAGACGCCAACAAGGTAGGAGCTGGAGCATTCGGGCTAGGGTTACCCCCACCGCAC 1214

Qy      181 GGAGGCCTTTTGGGGTGGAGCCCTCAGGCTCAGGGCATAACACAAACCTTGCCAGCAAAT 240
          |||
Db      1215 GGAGGCCTTTTGGGGTGGAGCCCTCAGGCTCAGGGCATAATGCAAACCTTGCCAGCAAAT 1274

Qy      241 CCGCCTCCTGCTTCCACCAATCGCCAGTCAGGAAGGCAGCCTACCCCGCTGTCTCCACCT 300
```



Db	1275	CCGCCTCCTGCCTCTACCAATCGCCAGTCAGGACGGCAGCCTACCCCGCTGTCTCCACCT	1334
Qy	301	TTGAGAAACACTCATCTCAAGCCATGCAGTGGAAGTCCACAACCTTTCCACCAAACCTCTG	360
Db	1335	CTGAGAACCACCTCATCTCAGGCCATGCACTGGAAGTCCACAACCTTCCACCAAACCTCTG	1394
Qy	361	CAAGATCCCAGAGTGAGAGGTCTGTATTTCCCTGCTGGTGGCTCCAGTTCAGGAACAGTA	420
Db	1395	CAAGATCCCAGAGTGAGAGGCCTGTATTTCCCTGCTGGTGGCTCCAGTTCAGGGACAGTA	1454
Qy	421	AACCCTGTTCCGACTACTGTCTCTCCCATATCGTCAATCTTCTCGAGGATTGGGGACCCCT	480
Db	1455	AACCCTGTTCCGACTACTACCTCTCCCATATCGTCAATCTTCTCGAGGATTGGGGACCCCT	1514
Qy	481	GCGCGGAACATGGAGAACATCACATCAGGATTCTTAGGACCCCTGCTCGTGTACAGGCG	540
Db	1515	GCGCTGAACATGGAGAACATCACATCAGGATTCTTAGGACCCCTGCTCGTGTACAGGCG	1574
Qy	541	GGGTTTTTCTTGTGACAAGAATCCTCACAAATACCGCAGAGTCTAGACTCGTGGTGGACT	600
Db	1575	GGGTTTTTCTTGTGACAAGAATCCTCACAAATACCGCAGAGTCTAGACTCGTGGTGGACT	1634
Qy	601	TCTCTCAATTTTCTAGGGGGAACACCGTGTGTCTTGGCCAAAATTTCGAGTCCCCAAC	660
Db	1635	TCTCTCAATTTTCTAGGGGGAACACCGTGTGTCTTGGCCAAAATTTCGAGTCCCCAATC	1694
Qy	661	TCCAATCACTACCAACCTCCTGTCTCCTCAACTTGTCTGTTATCGCTGGATGTGTCTG	720
Db	1695	TCCAATCACTACCAACCTCCTGTCTCCTCAACTTGTCTGTTATCGCTGGATGTGTCTG	1754
Qy	721	CGGCGTTTTATCATCTTCTCTTCATCCTGCTGCTATGCCTCATCTTCTTGTGGTTCTT	780
Db	1755	CGGCGTTTTATCATCTTCTCTTCATCCTGCTGCTATGCCTCATCTTCTTGTGGTTCTT	1814
Qy	781	CTGGACTATCAAGGTATGTTGCCCGTTTGTCTCTAATTCCAGGATCTTCAACCACCAGC	840
Db	1815	CTGGACTATCAAGGTATGTTGCCCGTTTGTCTCTAATTCCAGGATCATCAACCACCAGC	1874
Qy	841	ACGGGACCATGCAGAGCCTGCACGACTCCTGCTCAAGGAACCTCTATGTATCCCTCCTGT	900
Db	1875	ACGGGATCCTGCAGAACCTGCACGACTCCTGCTCAAGGAATCTCTATGTATCCCTCCTGT	1934
Qy	901	TGCTGTACAAAACCTTCGGATGGAACTGCACCTGTATTCCCATCCCATCATCCTGGGCT	960
Db	1935	TGCTGTACAAAACCTTCGGATGGAACTGCACCTGTATTCCCATCCCATCATCCTGGGCT	1994
Qy	961	TTCGGAAAATTCTATGGGAGTGGGCCTCAGCCCGTTTCTCCTGGCTCAGTTTACTAGTG	1020
Db	1995	TTCGGAAAATTCTATGGGAGTGGGCCTCAGCCCGTTTCTCCTGGCTCAGTTTACTAGTG	2054
Qy	1021	CCATTTGTTCAGTGGTTCGTAGGGCTTTCCCCACTGTTTGGCTTTTCAATTATATGGATG	1080
Db	2055	CCATTTGTTCAGTGGTTCGTAGGGCTTTCCCCATTGTTTGGCTTTTCAATTATATGGATG	2114
Qy	1081	ATGTTGTACTGGGGGCCAAGTCTGTACACCATCTTGAGTCCCTTTTTACCGCTGTTACCA	1140
Db	2115	ATGTTGTATGGGGGCCAAGTCTGTACAGCATCTTGAGTCCCTTTTTACCGCTGTTACCA	2174
Qy	1141	ATTTTCTTTTGTCTTTGGGTATACATTTAAACCTAATAAA	1181
Db	2175	ATTTTCTTTTGTCTTTGGGCATACATTTAAACCTAACA	2215



Qy	121	TGGCCAGACGCCAACAAAGGTAGGAGCTGGAGCATTCGGACTGGGGTTCACCCCACCGCAC	180
Db	1155	TGGCCAGACGCCAACAAAGGTAGGAGCTGGAGCATTCGGGCTAGGGTTCACCCCACCGCAC	1214
Qy	181	GGAGGCCTTTTGGGGTGGAGCCCTCAGGCTCAGGGCATAACACAAACCTTGCCAGCAAAT	240
Db	1215	GGAGGCCTTTTGGGGTGGAGCCCTCAGGCTCAGGGCATAATGCAAACCTTGCCAGCAAAT	1274
Qy	241	CCGCCTCCTGCTTCCACCAATCGCCAGTCAGGAAGGCAGCCTACCCCGCTGTCTCCACCT	300
Db	1275	CCGCCTCCTGCTTCTACCAATCGCCAGTCAGGACGGCAGCCTACCCCGCTGTCTCCACCT	1334
Qy	301	TTGAGAAACACTCATCCTCAAGCCATGCAGTGGAATCCACAACCTTCCACCAAACCTCTG	360
Db	1335	CTGAGAACCACTCATCCTCAGGCCATGCACTGGAATCCACAACCTTCCACCAAACCTCTG	1394
Qy	361	CAAGATCCAGAGTGAGAGGTCTGTATTTCCCTGCTGGTGGCTCCAGTTCAGGAACAGTA	420
Db	1395	CAAGATCCAGAGTGAGAGGCCTGTATTTCCCTGCTGGTGGCTCCAGTTCAGGGACAGTA	1454
Qy	421	AACCCTGTTCCGACTACTGTCTCTCCCATATCGTCAATCTTCTCGAGGATTGGGGACCCT	480
Db	1455	AACCCTGTTCCGACTACTACCTCTCCCATATCGTCAATCTTCTCGAGGATTGGGGACCCT	1514
Qy	481	GCGCGGAACATGGAGAACATCACATCAGGATTCTAGGACCCCTGCTCGTGTACAGGCG	540
Db	1515	GCGCTGAACATGGAGAACATCACATCAGGATTCTAGGACCCCTGCTCGTGTACAGGCG	1574
Qy	541	GGGTTTTTCTTGTGACAAGAATCCTCACAATACCGCAGAGTCTAGACTCGTGGTGGACT	600
Db	1575	GGGTTTTTCTTGTGACAAGAATCCTCACAATACCGCAGAGTCTAGACTCGTGGTGGACT	1634
Qy	601	TCTCTCAATTTTCTAGGGGGAACCTACCGTGTGTCTTGGCCAAAATTTCGCAGTCCCCAACC	660
Db	1635	TCTCTCAATTTTCTAGGGGGAACCTACCGTGTGTCTTGGCCAAAATTTCGCAGTCCCCAATC	1694
Qy	661	TCCAATCACTACCAACCTCCTGTCTCCAACCTTGTCTGGTTATCGCTGGATGTGTCTG	720
Db	1695	TCCAATCACTACCAACCTCCTGTCTCCAACCTTGTCTGGTTATCGCTGGATGTGTCTG	1754
Qy	721	CGGCGTTTTATCATCTTCTCTTCATCCTGCTGCTATGCCTCATCTTCTTGTGGTTCTT	780
Db	1755	CGGCGTTTTATCATCTTCTCTTCATCCTGCTGCTATGCCTCATCTTCTTGTGGTTCTT	1814
Qy	781	CTGGACTATCAAGGTATGTTGCCCGTTTGTCTCTAATTCCAGGATCTTCAACCACCAGC	840
Db	1815	CTGGACTATCAAGGTATGTTGCCCGTTTGTCTCTAATTCCAGGATCATCAACCACCAGC	1874
Qy	841	ACGGGACCATGCAGAGCCTGCACGACTCCTGCTCAAGGAACCTCTATGTATCCCTCCTGT	900
Db	1875	ACGGGATCCTGCAGAACCTGCACGACTCCTGCTCAAGGAATCTCTATGTATCCCTCCTGT	1934
Qy	901	TGCTGTACAAAACCTTCGGATGGAACTGCACCTGTATTCCCATCCCATCATCCTGGGCT	960
Db	1935	TGCTGTACAAAACCTTCGGATGGAACTGCACCTGTATTCCCATCCCATCATCCTGGGCT	1994
Qy	961	TTCGGAAAATTCTATGGGAGTGGGCCTCAGCCCGTTTCTCTGGCTCAGTTTACTAGTG	1020
Db	1995	TTCGGAAAATTCTATGGGAGTGGGCCTCAGCCCGTTTCTCTGGCTCAGTTTACTAGTG	2054
Qy	1021	CCATTTGTTTCAGTGGTTCGTAGGGCTTTCCCCCACTGTTTGGCTTTCAGTTATATGGATG	1080
Db	2055	CCATTTGTTTCAGTGGTTCGTAGGGCTTTCCCCCACTGTTTGGCTTTCAGTTATATGGATG	2114

```

Qy      1081 ATGTTGTACTGGGGGCCAAGTCTGTACACCATCTTGAGTCCCTTTTTACCGCTGTTACCA 1140
          ||||| ||| |||||||||||||||| ||||||||||||||||||||
Db      2115 ATGTGGTATTGGGGGCCAAGTCTGTACAGCATCTTGAGTCCCTTTTTACCGCTGTTACCA 2174

Qy      1141 ATTTTCTTTTGTCTTTGGGTATACATTAAACCCTAATAAA 1181
          |||||||||||||||||| |||||||||||||||||| |||
Db      2175 ATTTTCTTTTGTCTTTGGGCATACATTAAACCCTAACAAA 2215

```

## RESULT 5

HPBADYW

LOCUS HPBADYW 2743 bp DNA linear VRL 05-DEC-2005

DEFINITION human hepatitis b virus subtype adyw antigen genes (core antigen and surface antigen).

ACCESSION J02202

VERSION J02202.1 GI:329637

KEYWORDS .

SOURCE Hepatitis B virus

ORGANISM Hepatitis B virus

Viruses; Retro-transcribing viruses; Hepadnaviridae; Orthohepadnavirus.

REFERENCE 1 (bases 1 to 2743)

AUTHORS Pasek,M., Goto,T., Gilbert,W., Zink,B., Schaller,H., MacKay,P., Leadbetter,G. and Murray,K.

TITLE Hepatitis B virus genes and their expression in E. coli

JOURNAL Nature 282 (5739), 575-579 (1979)

PUBMED 399329

COMMENT Original source text: hbv subtype adyw from human.

cf hbvayw and whvsag. hbcag is hepatitis b core antigen protein, and hbsag is surface antigen protein.

## FEATURES

source Location/Qualifiers

```

1. .2743
   /organism="Hepatitis B virus"
   /mol_type="genomic DNA"
   /specific_host="Homo sapiens"
   /db_xref="taxon:10407"
   /note="subtype: adyw"

```

## CDS

```

88. .639
   /codon_start=1
   /product="core antigen"
   /protein_id="AAA45486.1"
   /db_xref="GI:329638"
   /translation="MDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTAAALYRDALES
PEHCSPHHTALRQAILCWGDLMTLATWVGNTNLEDPASRDLVVS YVNTNVGLKFRQLLW
FHISCLTFGRET VLEYLVSGVWIRTPPAYRPPNAPILSTLPETT VVRRRGRSPRRRT
PSPRRRRSQSPRRRRSQSRESQC"

```

## CDS

```

1524. .2204
   /codon_start=1
   /product="surface antigen"
   /protein_id="AAA45487.1"
   /db_xref="GI:329639"
   /translation="MENITSGFLGLP LLVLQAGFLLTRILTIPQSLDSWWTSLNFLGG
TTVCLGQNSQSPISNHSPTSCPPTCPGYRWMCLRRFII FLFILLCLIFLLVLLDYQG
MLPVCPLIPGSSTTSTGSCRTCTTPAQGISMYPSCCCTKPSDGNCTCIPSSWAFGK
FLWEWASARFSWLSLLVPFVQWVGLSPIVWLSVIWMMWYWGPSLYSILSPFLPLLPI
FFCLWAYI"

```

ORIGIN ?.

```

Query Match      95.9%;  Score 1133;  DB 10;  Length 2743;
Best Local Similarity  97.5%;  Pred. No. 0;
Matches 1151;  Conservative  0;  Mismatches  30;  Indels  0;  Gaps  0;

```

```

Qy      1 ATGGGGCAGAATCTTTCCACCAGCAATCCTCTGGGATTCTTTCCCGACCACCAGTTGGAT 60
          ||||||||||||||||||||||||||||||||||||||||||||

```

Db	1035	ATGGGGCAGAATCTTTCCACCAGCAATCCTCTGGGATTCTTTCCCGACCACCAGTTGGAT	1094
Qy	61	CCAGCCTTCAGAGCAAACACCAACAATCCAGATTGGGACTTCAATCCCAACAAGGACACC	120
Db	1095	CCAGCCTTCAGAGCAAACACCAACAATCCAGATTGGGACTTCAATCCCAACAAGGACACC	1154
Qy	121	TGGCCAGACGCCAACAAAGGTAGGAGCTGGAGCATTGGGACTGGGGTTCACCCCACCGCAC	180
Db	1155	TGGCCAGACGCCAACAAAGGTAGGAGCTGGAGCATTGGGGCTAGGGTTCACCCCACCGCAC	1214
Qy	181	GGAGGCCTTTTGGGGTGGAGCCCTCAGGCTCAGGGCATAACACAAACCTTGCCAGCAAAT	240
Db	1215	GGAGGCCTTTTGGGGTGGAGCCCTCAGGCTCAGGGCATAATGCAAACCTTGCCAGCAAAT	1274
Qy	241	CCGCCTCCTGCTTCCACCAATCGCCAGTCAGGAAGGCAGCCTACCCCGCTGTCTCCACCT	300
Db	1275	CCGCCTCCTGCTTCCACCAATCGCCAGTCAGGACGGCAGCCTACCCCGCTGTCTCCACCT	1334
Qy	301	TTGAGAAACACTCATCCTCAAGCCATGCAGTGGAACTCCACAACCTTCCACCAAACCTCTG	360
Db	1335	CTGAGAACCCTCATCCTCAGGCCATGCACTGGAACCTCCACAACCTTCCACCAAACCTCTG	1394
Qy	361	CAAGATCCCAGAGTGAGAGGTCTGTATTTCCCTGCTGGTGGCTCCAGTTCAGGAACAGTA	420
Db	1395	CAAGATCCCAGAGTGAGAGGCCTGTATTTCCCTGCTGGTGGCTCCAGTTCAGGGACAGTA	1454
Qy	421	AACCCTGTTCCGACTACTGTCTCTCCCATATCGTCAATCTTCTCGAGGATTGGGGACCCT	480
Db	1455	AACCCTGTTCCGACTACTACCTCTCCCATATCGTCAATCTTCTCGAGGATTGGGGACCCT	1514
Qy	481	GCGCGGAACATGGAGAACATCACATCAGGATTCTAGGACCCCTGCTCGTGTTACAGGCG	540
Db	1515	GCGCTGAACATGGAGAACATCACATCAGGATTCTAGGACCCCTGCTCGTGTTACAGGCG	1574
Qy	541	GGGTTTTTCTTGTTGACAAGAATCCTCACAATACCGCAGAGTCTAGACTCGTGGTGGACT	600
Db	1575	GGGTTTTTCTTGTTGACAAGAATCCTCACAATACCGCAGAGTCTAGACTCGTGGTGGACT	1634
Qy	601	TCTCTCAATTTTCTAGGGGGAACACCGTGTGTCTTGGCCAAAATTTCGCAGTCCCCAACC	660
Db	1635	TCTCTCAATTTTCTAGGGGGAACACCGTGTGTCTTGGCCAAAATTTCGCAGTCCCCAATC	1694
Qy	661	TCCAATCACTACCAACCTCCTGTCTCCTCAACTTGTCTGGTTATCGCTGGATGTGTCTG	720
Db	1695	TCCAATCACTACCAACCTCCTGTCTCCTCAACTTGTCTGGTTATCGCTGGATGTGTCTG	1754
Qy	721	CGGCGTTTTATCATCTTCTCTTCATCCTGCTGCTATGCCTCATCTTCTTGTGGTTCTT	780
Db	1755	CGGCGTTTTATCATCTTCTCTTCATCCTGCTGCTATGCCTCATCTTCTTGTGGTTCTT	1814
Qy	781	CTGGACTATCAAGGTATGTTGCCCGTTTGTCTCTAATTCCAGGATCTTCAACCACCAGC	840
Db	1815	CTGGACTATCAAGGTATGTTGCCCGTTTGTCTCTAATTCCAGGATCATCAACCACCAGC	1874
Qy	841	ACGGGACCATGCAGAGCCTGCACGACTCCTGCTCAAGGAACCTCTATGTATCCCTCCTGT	900
Db	1875	ACGGGATCCTGCAGAACCTGCACGACTCCTGCTCAAGGAATCTCTATGTATCCCTCCTGT	1934
Qy	901	TGCTGTACAAAACCTTCGGATGGAACTGCACCTGTATTCCCATCCCATCATCTGGGCT	960
Db	1935	TGCTGTACAAAACCTTCGGATGGAACTGCACCTGTATTCCCATCCCATCATCTGGGCT	1994
Qy	961	TTCGGAAAATTCCTATGGGAGTGGGCCTCAGCCCGTTTCTCCTGGCTCAGTTTACTAGTG	1020

```

Db      1995  TTCGGAATTCCTATGGGAGTGGGCCTCAGCCCGTTTCTCTTGGCTCAGTTTACTAGTG 2054
Qy      1021  CCATTTGTTTCAGTGGTTCGTAGGGCTTTCCCCACTGTTTGGCTTTCAGTTATATGGATG 1080
          ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      2055  CCATTTGTTTCAGTGGTTCGTAGGGCTTTCCCCATTGTTTGGCTTTCAGTTATATGGATG 2114
Qy      1081  ATGTTGTACTGGGGCCAAGTCTGTACACCATCTTGAGTCCCTTTTTACCGCTGTTACCA 1140
          |||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      2115  ATGTGGTATTGGGGCCAAGTCTGTACAGCATCTTGAGTCCCTTTTTACCGCTGTTACCA 2174
Qy      1141  ATTTTCTTTTGTCTTTGGGTATACATTTAAACCTAATAAA 1181
          ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      2175  ATTTTCTTTTGTCTTTGGGCATACATTTAAACCTAACAAA 2215

```

## RESULT 6

AB104723

LOCUS AB104723 1170 bp DNA linear VRL 26-JUN-2003

DEFINITION Hépatitis B virus s gene for pre-S and S protein, complete cds,  
isolate: EG91.

ACCESSION AB104723

VERSION AB104723.1 GI:32261206

KEYWORDS .

SOURCE Hepatitis B virus

ORGANISM Hepatitis B virus  
Viruses; Retro-transcribing viruses; Hepadnaviridae;  
Orthohepadnavirus.

REFERENCE 1

AUTHORS Saudy,N., Sugauchi,F., Tanaka,Y., Suzuki,S., Aal,A.A., Zaid,M.A.,  
Agha,S. and Mizokami,M.TITLE Genotypes and phylogenetic characterization of hepatitis B and  
delta viruses in Egypt

JOURNAL J. Med. Virol. 70 (4), 529-536 (2003)

PUBMED 12794714

REFERENCE 2 (bases 1 to 1170)

AUTHORS Suzuki,S., Saudy,N., Sugauchi,F., Orito,E., Agha,S. and Mizokami,M.

TITLE Direct Submission

JOURNAL Submitted (26-FEB-2003) Seiji Suzuki, Nagoya City University  
Graduate School, Department of clinical Molecular Informative  
Medicine; Mizuho, Nagoya, Aichi 467-8601, Japan  
(E-mail:seijis@med.nagoya-cu.ac.jp, Tel:81-52-853-8292,  
Fax:81-52-842-0021)

FEATURES Location/Qualifiers

```

source      1. .1170
              /organism="Hepatitis B virus"
              /mol_type="genomic DNA"
              /isolate="EG91"
              /db_xref="taxon:10407"
              /country="Egypt"
              /note="genotype D"

```

```

gene        1. .1170
              /gene="s"

```

```

CDS          1. .1170
              /gene="s"
              /codon_start=1
              /product="pre-S and S protein"
              /protein_id="BAC78518.1"
              /db_xref="GI:32261207"
              /translation="MGQNLSTSNPLGFFPDHQLDPAFRANTANPDWDFNPNKDTWPDAN
              KVGAGAFGLGFTPPHGGLLGWSPQAQGILQTLPANPPPASTNRQSGRQPTPLSPPLR
              NTHPQAMQWNSTTFHQTLDPRVRGLYFPAGGSSSGTVSPVPTTVSHISSIFSRIGDP
              ALNMENITSGFLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGGTTVCLGQNSQS
              PTSNHSPTSCPPTCPGYRWMCLRRFIIIFLIFLLVLLDYQGMLPVCPLIPGS
              STTSTGPCRTCTTPAQGTSMPSCCCTKPLDGNCTCIPISSWAFGKFLWEWASARFS

```

WLSLLVPFVOWFVGLSPTVWLSVIWMMWYWGPSLYSILSPFLPLLPIFFCLWVYI"

## ORIGIN

Query Match 95.8%; Score 1131.6; DB 10; Length 1170;  
Best Local Similarity 97.9%; Pred. No. 0;  
Matches 1146; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

Qy	1	ATGGGGCAGAATCTTTCCACCAGCAATCCTCTGGGATTCTTTCCCGACCACCAGTTGGAT	60
Db	1	ATGGGGCAGAATCTTTCCACCAGCAATCCTCTGGGATTCTTTCCCGACCACCAGTTGGAT	60
Qy	61	CCAGCCTTCAGAGCAAACACCAACAATCCAGATTGGGACTTCAATCCCAACAAGGACACC	120
Db	61	CCAGCCTTCAGAGCAAACACCGCAAATCCAGATTGGGACTTCAATCCCAACAAGGACACC	120
Qy	121	TGGCCAGACGCCAACAAAGGTAGGAGCTGGAGCATTTCGGACTGGGGTTCACCCACCGCAC	180
Db	121	TGGCCAGACGCCAACAAAGGTAGGAGCTGGAGCATTTCGGGCTGGGATTTCACCCACCGCAC	180
Qy	181	GGAGGCCTTTTGGGGTGGAGCCCTCAGGCTCAGGGCATAACACAAACCTTGCCAGCAAAT	240
Db	181	GGAGGCCTTTTGGGGTGGAGCCCTCAGGCTCAGGGCATACTACAAACCTTGCCAGCAAAT	240
Qy	241	CCGCCTCCTGCTTCCACCAATCGCCAGTCAGGAAGGCAGCCTACCCCGCTGTCTCCACCT	300
Db	241	CCGCCTCCTGCTTCTACCAATCGCCAGTCAGGAAGGCAGCCTACCCCTCTGTCTCCACCT	300
Qy	301	TTGAGAAACACTCATCCTCAAGCCATGCAGTGGAACTCCACAACCTTCCACCAAACCTCTG	360
Db	301	CTGAGAAACACTCATCCTCAGGCCATGCAGTGGAACTCCACAACCTTCCACCAAACCTCTG	360
Qy	361	CAAGATCCCAGAGTGAGAGGTCTGTATTTCCCTGCTGGTGGCTCCAGTTCAGGAACAGTA	420
Db	361	CAAGATCCCAGAGTGAGAGGCCTGTATTTCCCTGCTGGTGGCTCCAGTTCAGGAACAGTA	420
Qy	421	AACCCTGTTCCGACTACTGTCTCTCCCATATCGTCAATCTTCTCGAGGATTGGGGACCT	480
Db	421	AGCCCTGTTCCGACTACTGTCTCTCACATATCGTCAATCTTCTCGAGGATTGGGGACCT	480
Qy	481	GCGCGGAACATGGAGAACATCACATCAGGATTCCCTAGGACCCCTGCTCGTGTTACAGGCG	540
Db	481	GCGCTGAACATGGAGAACATCACATCAGGATTCCCTAGGACCCCTGCTCGTGTTACAGGCG	540
Qy	541	GGGTTTTTCTTGTGACAAGAATCCTCACAATACCGCAGAGTCTAGACTCGTGGTGGACT	600
Db	541	GGGTTTTTCTTGTGACAAGAATCCTCACAATACCGCAGAGTCTAGACTCGTGGTGGACT	600
Qy	601	TCTCTCAATTTTCTAGGGGGAACTACCGTGTGTCTTGGCCAAAATTTCGAGTCCCCAACC	660
Db	601	TCTCTCAATTTTCTAGGGGGAACTACCGTGTGTCTTGGCCAAAATTTCGAGTCCCCAACC	660
Qy	661	TCCAATCACTACCAACCTCCTGTCCTCCAACCTTGTCCTGGTTATCGCTGGATGTGTCTG	720
Db	661	TCCAATCACTACCAACCTCCTGTCCTCCAACCTTGTCCTGGTTATCGCTGGATGTGTCTG	720
Qy	721	CGGCGTTTTATCATCTTCCTCTTCATCCTGCTGCTATGCCTCATCTTCTTGTGGTTCTT	780
Db	721	CGGCGTTTTATCATCTTCCTCTTCATCCTGCTGCTATGCCTCATCTTCTTGTGGTTCTT	780
Qy	781	CTGGACTATCAAGGTATGTTGCCCCGTTTGTCTCTAATTCAGGATCTTCAACCACCAGC	840
Db	781	CTGGACTATCAAGGTATGTTGCCCCGTTTGTCTCTAATTCAGGATCTTCAACCACCAGC	840
Qy	841	ACGGGACCATGCAGAGCCTGCACGACTCCTGCTCAAGGAACCTCTATGTATCCCTCCTGT	900

```

      |||
Db      841  ACGGGACCATGCAGAACCTGCACGACTCCTGCTCAAGGAACCTCTATGTATCCCTCCTGT 900
      |||
Qy      901  TGCTGTACAAACCTTCGGATGGAACTGCACCTGTATTCCCATCCCATCATCCTGGGCT 960
      |||
Db      901  TGCTGTACAAACCTTTGGACGGAAATTGCACCTGTATTCCCATCCCATCATCCTGGGCT 960
      |||
Qy      961  TTCGGAAAATTCTATGGGAGTGGGCCTCAGCCCGTTTCTCCTGGCTCAGTTTACTAGTG 1020
      |||
Db      961  TTCGGAAAATTCTATGGGAGTGGGCCTCAGCCCGTTTCTCCTGGCTCAGTTTACTAGTG 1020
      |||
Qy      1021 CCATTTGTTTCAGTGGTTCGTAGGGCTTTCCCCCACTGTTTGGCTTTCAGTTATATGGATG 1080
      |||
Db      1021 CCATTTGTTTCAGTGGTTCGTAGGGCTTTCCCCCACTGTTTGGCTTTCAGTTATATGGATG 1080
      |||
Qy      1081 ATGTTGTACTGGGGGCCAAGTCTGTACACCATCTTGAGTCCCTTTTTACCGCTGTTACCA 1140
      |||
Db      1081 ATGTTGTATTGGGGGCCAAGTCTGTACAGCATCTTGAGTCCCTTTTTACCGCTGTTACCA 1140
      |||
Qy      1141 ATTTTCTTTTGTCTTTGGGTATACATTTAA 1170
      |||
Db      1141 ATTTTCTTTTGTCTTTGGGTATACATTTAA 1170

```

## RESULT 7

AB104715

LOCUS AB104715 1170 bp DNA linear VRL 26-JUN-2003

DEFINITION Hepatitis B virus s gene for pre-S and S protein, complete cds,  
isolate: EG69.

ACCESSION AB104715

VERSION AB104715.1 GI:32261190

KEYWORDS .

SOURCE Hepatitis B virus

ORGANISM Hepatitis B virus  
Viruses; Retro-transcribing viruses; Hepadnaviridae;  
Orthohepadnavirus.

REFERENCE 1

AUTHORS Saudy,N., Sugauchi,F., Tanaka,Y., Suzuki,S., Aal,A.A., Zaid,M.A.,  
Agha,S. and Mizokami,M.TITLE Genotypes and phylogenetic characterization of hepatitis B and  
delta viruses in Egypt

JOURNAL J. Med. Virol. 70 (4), 529-536 (2003)

PUBMED 12794714

REFERENCE 2 (bases 1 to 1170)

AUTHORS Suzuki,S., Saudy,N., Sugauchi,F., Orito,E., Agha,S. and Mizokami,M.

TITLE Direct Submission

JOURNAL Submitted (26-FEB-2003) Seiji Suzuki, Nagoya City University  
Graduate School, Department of clinical Molecular Informative  
Medicine; Mizuho, Nagoya, Aichi 467-8601, Japan  
(E-mail:seijis@med.nagoya-cu.ac.jp, Tel:81-52-853-8292,  
Fax:81-52-842-0021)

FEATURES Location/Qualifiers

```

source      1. .1170
              /organism="Hepatitis B virus"
              /mol_type="genomic DNA"
              /isolate="EG69"
              /db_xref="taxon:10407"
              /country="Egypt"
              /note="genotype D"

gene        1. .1170
              /gene="s"

CDS         1. .1170
              /gene="s"
              /codon_start=1

```



```

/product="pre-S and S protein"
/protein_id="BAC78510.1"
/db_xref="GI:32261191"
/translation="MGQNLSTSNPLGFFPDHQLLPAPFRANTANPDWDFNPKNKDTWPDAN
KVGAGAFGLGFTPPHGGLLGWSPOAQGILQTLPANPPPASTNRQSGRQPTPLSPPLR
NTHPQAMQWNSTTFHQTLQDPRVRGLYFPAGGSSSGTVNPVPTTVSPISSIFSRIGDP
ALNMENITSGFLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGGTTVCLGQNSQS
PTSNHSPTSCPPTCPGYRWMCLRRFIIIFLIFLLCLIFLLVLLDYQGMPLVCPCLIPGS
STTSTGPCRTCTTPAQGTSMPSCCCTKPLDGNCTCIPISSWAFGKFLWEWASARFS
WLSLLVPFVQWFVGLSPTVWLSVIWMMWYWGPSLYSILSPFLPLLPIFFCLWVYI"

```

ORIGIN

Query Match	95.5%;	Score 1128.4;	DB 10;	Length 1170;
Best Local Similarity	97.8%;	Pred. No. 0;		
Matches 1144;	Conservative 0;	Mismatches 26;	Indels 0;	Gaps 0;
Qy	1	ATGGGGCAGAATCTTTCCACCAGCAATCCTCTGGGATTCTTTCCCGACCACCAGTTGGAT	60	
Db	1	ATGGGGCAGAATCTTTCCACCAGCAATCCTCTGGGATTCTTTCCCGACCACCAGTTGTTA	60	
Qy	61	CCAGCCTTCAGAGCAAACACCAACAATCCAGATTGGGACTTCAATCCCAACAAGGACACC	120	
Db	61	CCAGCCTTCAGAGCAAACACCGCAAATCCAGATTGGGACTTCAATCCCAACAAGGACACC	120	
Qy	121	TGGCCAGACGCCAACAAAGGTAGGAGCTGGAGCATTCGGACTGGGGTTCACCCCACCGCAC	180	
Db	121	TGGCCAGACGCCAACAAAGGTAGGAGCTGGAGCATTCGGGGCTGGGATTCACCCCACCGCAC	180	
Qy	181	GGAGGCCTTTTGGGGTGGAGCCCTCAGGCTCAGGGCATAACACAAACCTTGCCAGCAAAT	240	
Db	181	GGAGGCCTTTTGGGGTGGAGCCCTCAGGCTCAGGGCATACTACAAACCTTGCCAGCAAAT	240	
Qy	241	CCGCCTCCTGCTTCCACCAATCGCCAGTCAGGAAGGCAGCCTACCCCGCTGTCTCCACCT	300	
Db	241	CCGCCTCCTGCTTCTACCAATCGCCAGTCAGGAAGGCAGCCTACCCCTCTGTCTCCACCT	300	
Qy	301	TTGAGAAACACTCATCCTCAAGCCATGCAGTGGAACCTCCACAACCTTCCACCAAACCTCTG	360	
Db	301	CTGAGAAACACTCATCCTCAGGCCATGCAGTGGAACCTCCACAACCTTCCACCAAACCTCTG	360	
Qy	361	CAAGATCCCAGAGTGAGAGGTCTGTATTTCCCTGCTGGTGGCTCCAGTTCAGGAACAGTA	420	
Db	361	CAAGATCCCAGAGTGAGAGGCCTGTATTTCCCTGCTGGTGGCTCCAGTTCAGGAACAGTA	420	
Qy	421	AACCCTGTTCCGACTACTGTCTCTCCCATATCGTCAATCTTCTCGAGGATTGGGGACCCT	480	
Db	421	AACCCTGTTCCGACTACTGTCTCTCCCATATCGTCAATCTTCTCGAGGATTGGGGACCCT	480	
Qy	481	GCGCGGAACATGGAGAACATCACATCAGGATTCTAGGACCCCTGCTCGTGTACAGGCG	540	
Db	481	GCGTTGAACATGGAGAACATCACATCAGGATTCTAGGACCCCTGCTCGTGTACAGGCG	540	
Qy	541	GGGTTTTTCTTGTTGACAAGAATCCTCACAATACCGCAGAGTCTAGACTCGTGGTGGACT	600	
Db	541	GGGTTTTTCTTGTTGACAAGAATCCTCACAATACCGCAGAGTCTAGACTCGTGGTGGACT	600	
Qy	601	TCTCTCAATTTTCTAGGGGAACTACCGTGTGTCTTGGCCAAAATTTCGCAGTCCCCAACC	660	
Db	601	TCTCTCAATTTTCTAGGGGAACTACCGTGTGTCTTGGCCAAAATTTCGCAGTCCCCAACC	660	
Qy	661	TCCAATCACTACCAACCTCCTGTCCTCCAACCTTGTCTGGTTATCGCTGGATGTGTCTG	720	
Db	661	TCCAATCACTACCAACCTCCTGTCCTCCAACCTTGTCTGGTTATCGCTGGATGTGTCTG	720	

```
Qy      721  CGGCGTTTTATCATCTTCCTCTTCATCCTGCTGCTATGCCTCATCTTCTTGTGGTTCTT 780
      |||
Db      721  CGGCGTTTTATCATCTTCCTCTTCATCCTGCTGCTATGCCTCATCTTCTTGTGGTTCTT 780

Qy      781  CTGGACTATCAAGGTATGTTGCCCGTTTGTCTCTAATTCCAGGATCTTCAACCACCAGC 840
      |||
Db      781  CTGGACTATCAAGGTATGTTGCCCGTTTGTCTCTAATTCCAGGATCTTCAACCACCAGC 840

Qy      841  ACGGGACCATGCAGAGCCTGCACGACTCCTGCTCAAGGAACCTCTATGTATCCCTCCTGT 900
      |||
Db      841  ACGGGACCATGCAGAACCTGCACGACTCCTGCTCAAGGAACCTCTATGTATCCCTCCTGT 900

Qy      901  TGCTGTACAAAACCTTCGGATGGAACTGCACCTGTATTCCCATCCCATCATCCTGGGCT 960
      |||
Db      901  TGCTGTACAAAACCTTTGGACGGAAATGCACCTGTATTCCCATCCCATCATCCTGGGCT 960

Qy      961  TTCGGAAAATTCTATGGGAGTGGGCCTCAGCCCGTTTCTCCTGGCTCAGTTTACTAGTG 1020
      |||
Db      961  TTCGGAAAATTCTATGGGAGTGGGCCTCAGCCCGTTTCTCCTGGCTCAGTTTACTAGTG 1020

Qy      1021 CCATTTGTTTCAGTGGTTCGTAGGGCTTCCCCCACTGTTTGGCTTTCAGTTATATGGATG 1080
      |||
Db      1021 CCATTTGTTTCAGTGGTTCGTAGGGCTTCCCCCACTGTTTGGCTTTCAGTTATATGGATG 1080

Qy      1081 ATGTTGTACTGGGGGCCAAGTCTGTACACCATCTTGAGTCCCTTTTTACCGCTGTTACCA 1140
      |||
Db      1081 ATGTGGTATTGGGGGCCAAGTCTGTACAGCATCTTGAGTCCCTTTTTACCGCTGTTACCA 1140

Qy      1141 ATTTTCTTTTGTCTTTGGGTATACATTTAA 1170
      |||
Db      1141 ATTTTCTTTTGTCTTTGGGTATACATTTAA 1170
```

## RESULT 8

BD181818

LOCUS BD181818 8007 bp DNA linear PAT 15-MAY-2003

DEFINITION Hepatitis B virus vectors for gene therapy.

ACCESSION BD181818

VERSION BD181818.1 GI:30792736

KEYWORDS JP 2002320480-A/3.

SOURCE unidentified

ORGANISM unidentified

unclassified sequences.

REFERENCE 1 (bases 1 to 8007)

AUTHORS Ryu,W., Jeong,J.K., Lee,J., Cho,W.Y. and Yoon,G.S.

TITLE Hepatitis B virus vectors for gene therapy

JOURNAL Patent: JP 2002320480-A 3 05-NOV-2002;

WANG-SHICK RYU

COMMENT OS Unidentified

PN JP 2002320480-A/3

PD 05-NOV-2002

PF 20-APR-2001 JP 2001122392

PR 20-APR-2000 KR 2000-21070,12-APR-2001 KR 2001-19645 PI

WANG-SHICK RYU,JONG KEUN JEONG,JEHAN LEE,WOO YOUNG CHO,GYE PI

SOON YOON

PC C12N15/09,A61K35/74,A61K48/00,A61P31/20,C12N7/00,C12N15/00 CC

pCMV-HBV/30 Full Sequences

CC 8007 bp ms-DNA circular

CC From HBV-ayw

CC #1 ; #1820 of HBV-ayw (accession number J02203) CC

transcription start site of HBV pregenomic RNA CC #1 - 3360 ;

HBV-ayw (178 bp overlapping the HBV genome-length) CC 5'- epsilon

secondary structure ; bases 30 - 90 CC 3'- epsilon secondary

structure ; bases 3212 - 3272 CC 5'- DR1 ; bases 7 - 17

CC 3'- DR1 ; bases 3189 - 3199  
 CC 3'- DR2 ; bases 2955 - 2965  
 CC Poly A signal ; bases 3281-3286  
 CC Core ORF ; bases 84 - 632 (exclude stop codon) CC Polymerase  
 ORF ; bases 490 - 2985 (exclude stop codon) CC S1 ORF ; bases  
 1031 - 2197 (exclude stop codon) CC S2 ORF ; bases 1355 - 2197  
 (exclude stop codon) CC S ORF ; bases 1520 - 2197 (exclude stop  
 codon) CC X ORF ; bases 2739 - 3200 (exclude stop codon) CC  
 From pcDNA1/Amp  
 CC Col E1 origin ; bases 5103-5689 (1-587 of pcDNA1/Amp) CC M13  
 origin ; bases 5690-6282  
 CC Ampicillin gene ; bases 6462-7405  
 CC CMV promoter ; bases 7406-7999  
 CC SP6 primer sequence ; bases 3372-3390  
 CC Splice and polyA ; bases 3391-4089  
 FH Key Location/Qualifiers  
 FT source 1. .8007  
 FT /organism='Unidentified'.

FEATURES Location/Qualifiers  
 source 1. .8007  
 /organism="unidentified"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:32644"

## ORIGIN

Query Match 95.5%; Score 1128.2; DB 2; Length 8007;  
 Best Local Similarity 97.2%; Pred. No. 0;  
 Matches 1148; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

Qy	1	ATGGGGCAGAATCTTTCCACCAGCAATCCTCTGGGATTCTTTCCCGACCACCAGTTGGAT	60
Db	1031	ATGGGGCAGAATCTTTCCACCAGCAATCCTCTGGGATTCTTTCCCGACCACCAGTTGGAT	1090
Qy	61	CCAGCCTTCAGAGCAAACACCAACAATCCAGATTGGGACTTCAATCCCAACAAGGACACC	120
Db	1091	CCAGCCTTCAGAGCAAACACCGCAAATCCAGATTGGGACTTCAATCCCAACAAGGACACC	1150
Qy	121	TGGCCAGACGCCAACAAGGTAGGAGCTGGAGCATTCGGACTGGGGTTCACCCCACCGCAC	180
Db	1151	TGGCCAGACGCCAACAAGGTAGGAGCTGGAGCATTCGGGCTGGGTTTCACCCCACCGCAC	1210
Qy	181	GGAGGCCTTTTGGGGTGGAGCCCTCAGGCTCAGGGCATAACACAAACCTTGCCAGCAAAT	240
Db	1211	GGAGGCCTTTTGGGGTGGAGCCCTCAGGCTCAGGGCATACTACAAACTTTGCCAGCAAAT	1270
Qy	241	CCGCCTCCTGCTTCCACCAATCGCCAGTCAGGAAGGCAGCCTACCCCGCTGTCTCCACCT	300
Db	1271	CCGCCTCCTGCTTCCACCAATCGCCAGTCAGGAAGGCAGCCTACCCCGCTGTCTCCACCT	1330
Qy	301	TTGAGAAACACTCATCCTCAAGCCATGCAGTGGAATCCACAACCTTCCACCAAACCTCTG	360
Db	1331	TTGAGAAACACTCATCCTCAGGCCATGCAGTGGAATCCACAACCTTCCACCAAACCTCTG	1390
Qy	361	CAAGATCCCAGAGTGAGAGGTCTGTATTTCCCTGCTGGTGGCTCCAGTTCAGGAACAGTA	420
Db	1391	CAAGATCCCAGAGTGAGAGGCCTGTATTTCCCTGCTGGTGGCTCCAGTTCAGGAACAGTA	1450
Qy	421	AACCCTGTTCCGACTACTGTCTCTCCCATATCGTCAATCTTCTCGAGGATTGGGGACCCT	480
Db	1451	AACCCTGTTCTGACTACTGCCTCTCCCTTATCGTCAATCTTCTCGAGGATTGGGGACCCT	1510
Qy	481	GCGCGGAACATGGAGAACATCACATCAGGATTCCTAGGACCCCTGCTCGTGTTACAGGCG	540
Db	1511	GCGCTGAACATGGAGAACATCACATCAGGATTCCTAGGACCCCTTCTCGTGTTACAGGCG	1570

```

Qy      541 GGGTTTTTCTTGTGACAAGAATCCTCACAATACCGCAGAGTCTAGACTCGTGGTGGACT 600
      |||
Db      1571 GGGTTTTTCTTGTGACAAGAATCCTCACAATACCGCAGAGTCTAGACTCGTGGTGGACT 1630

Qy      601 TCTCTCAATTTTCTAGGGGGAACACCGTGTGTCTTGGCCAAAATTTCGCAGTCCCCAACC 660
      |||
Db      1631 TCTCTCAATTTTCTAGGGGGAACACCGTGTGTCTTGGCCAAAATTTCGCAGTCCCCAACC 1690

Qy      661 TCCAATCACTACCAACCTCCTGTCCTCCAACCTTGTCTGGTTATCGCTGGATGTGTCTG 720
      |||
Db      1691 TCCAATCACTACCAACCTCCTGTCCTCCAACCTTGTCTGGTTATCGCTGGATGTGTCTG 1750

Qy      721 CGGCGTTTTATCATCTTCCTCTTCATCCTGCTGCTATGCCTCATCTTCTTGTGGTTCTT 780
      |||
Db      1751 CGGCGTTTTATCATCTTCCTCTTCATCCTGCTGCTATGCCTCATCTTCTTGTGGTTCTT 1810

Qy      781 CTGGACTATCAAGGTATGTTGCCCGTTTGTCTCTAATTCCAGGATCTTCAACCACCAGC 840
      |||
Db      1811 CTGGACTATCAAGGTATGTTGCCCGTTTGTCTCTAATTCCAGGATCTTCAACAACCAGC 1870

Qy      841 ACGGGACCATGCAGAGCCTGCACGACTCCTGCTCAAGGAACCTCTATGTATCCCTCCTGT 900
      |||
Db      1871 ACGGGACCATGCCGACCTGCATGACTACTGCTCAAGGAACCTCTATGTATCCCTCCTGT 1930

Qy      901 TGCTGTACAAAACCTTCGGATGGAAACTGCACCTGTATTCCCATCCCATCATCCTGGGCT 960
      |||
Db      1931 TGCTGTACAAAACCTTCGGACGAAATTGCACCTGTATTCCCATCCCATCATCCTGGGCT 1990

Qy      961 TTCGGAAAATTCCTATGGGAGTGGGCCTCAGCCCGTTTCTCCTGGCTCAGTTTACTAGTG 1020
      |||
Db      1991 TTCGGAAAATTCCTATGGGAGTGGGCCTCAGCCCGTTTCTCCTGGCTCAGTTTACTAGTG 2050

Qy      1021 CCATTTGTTTCAGTGGTTCGTAGGGCTTTCCCCCACTGTTTGGCTTTCAGTTATATGGATG 1080
      |||
Db      2051 CCATTTGTTTCAGTGGTTCGTAGGGCTTTCCCCCACTGTTTGGCTTTCAGTTATATGGATG 2110

Qy      1081 ATGTTGTACTGGGGGCCAAGTCTGTACACCATCTTGAGTCCCTTTTTACCGCTGTTACCA 1140
      |||
Db      2111 ATGTGGTATTGGGGGCCAAGTCTGTACAGCATCTTGAGTCCCTTTTTACCGCTGTTACCA 2170

Qy      1141 ATTTTCTTTTGTCTTTGGGTATACATTTAAACCCTAATAAA 1181
      |||
Db      2171 ATTTTCTTTTGTCTTTGGGTATACATTTAAACCCTAACAAA 2211

```

## RESULT 9

AY603461

LOCUS AY603461 1170 bp DNA linear VRL 19-APR-2005  
 DEFINITION Hepatitis B virus isolate 04T large S protein (S), middle S protein (S), and S protein (S) genes, complete cds.

ACCESSION AY603461

VERSION AY603461.1 GI:47499935

KEYWORDS

SOURCE Hepatitis B virus

ORGANISM Hepatitis B virus

Viruses; Retro-transcribing viruses; Hepadnaviridae;  
 Orthohepadnavirus.

REFERENCE 1 (bases 1 to 1170)

AUTHORS Sominskaya,I., Mihailova,M., Jansons,J., Emelyanova,V.,  
 Folkmane,I., Smagris,E., Dumpis,U., Rozentals,R. and Pumpens,P.

TITLE Hepatitis B and C virus variants in long-term immunosuppressed  
 renal transplant patients in Latvia

JOURNAL Intervirology 48 (2-3), 192-200 (2005)

PUBMED 15812194  
 REFERENCE 2 (bases 1 to 1170)  
 AUTHORS Sominskaya,I., Mihailova,M., Jansons,J., Emelyanova,V.,  
 Folkmane,I., Smagris,E., Dumpis,U., Rozentals,R. and Pumpens,P.  
 TITLE Direct Submission  
 JOURNAL Submitted (21-APR-2004) Protein Engineering Department, Biomedical  
 Research and Study Centre, University of Latvia, Ratsupites str. 1,  
 Riga LV-1067, Latvia

FEATURES Location/Qualifiers  
 source 1. .1170  
 /organism="Hepatitis B virus"  
 /virion  
 /mol\_type="genomic DNA"  
 /isolate="04T"  
 /isolation\_source="renal transplantation patient"  
 /db\_xref="taxon:10407"  
 /country="Latvia"  
 /note="subtype: ayw3; genotype: D"  
 gene 1. .1170  
 /gene="S"  
 CDS 1. .1170  
 /gene="S"  
 /note="pre-S1/pre-S2/S; LHBS/MHBS/HBsAg; surface antigen"  
 /codon\_start=1  
 /product="large S protein"  
 /protein\_id="AAT28720.1"  
 /db\_xref="GI:47499936"  
 /translation="MGQNLSTSNPLGFFPDHQLDPAFRANTANPDWDFNPNKDTWPDAN  
 KVGAGAGFGLGFTPPHGGLLGWSPQAQGI IQTL PANPPPASTNRQSGRQPTPLSPPLR  
 NTHPQAMQWNSTTFHQTLQDPRVRGLYFPAGGSSSGTVNPVPTTVSPISSIFSRIGDP  
 ALNMENITSGFLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGGTTVCLGQNSQS  
 PTSNHSPTSCPPTCPGYRWMCLRRFIIIFLIFLLVLLDYQGMLPVCPLIPGS  
 STTSAGPCRTCTTTAQGTSMPSCCCTKPSDGNCTCIPIPSSWAFGKFLWEWASARFS  
 WLSLLVPFVQWVGLSPTVWLSVIWMMWYGPSLYSILSPFLPLLPIFFCLWVYI"  
 CDS 325. .1170  
 /gene="S"  
 /note="pre-S2/S; MHBS/HBsAg; surface antigen"  
 /codon\_start=1  
 /product="middle S protein"  
 /protein\_id="AAT28721.1"  
 /db\_xref="GI:47499937"  
 /translation="MQWNSTTFHQTLQDPRVRGLYFPAGGSSSGTVNPVPTTVSPISS  
 IFSRIGDPALNMENITSGFLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGGTTV  
 CLGQNSQSPTSNSPTSCPPTCPGYRWMCLRRFIIIFLIFLLVLLDYQGMLP  
 VCPLIPGSSTTSAGPCRTCTTTAQGTSMPSCCCTKPSDGNCTCIPIPSSWAFGKFLW  
 EWASARFSWLSLLVPFVQWVGLSPTVWLSVIWMMWYGPSLYSILSPFLPLLPIFFC  
 LWVYI"  
 CDS 490. .1170  
 /gene="S"  
 /note="HBsAg; surface antigen"  
 /codon\_start=1  
 /product="S protein"  
 /protein\_id="AAT28722.1"  
 /db\_xref="GI:47499938"  
 /translation="MENITSGFLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGG  
 TTVCLGQNSQSPTSNSPTSCPPTCPGYRWMCLRRFIIIFLIFLLVLLDYQG  
 MLPVCPLIPGSSTTSAGPCRTCTTTAQGTSMPSCCCTKPSDGNCTCIPIPSSWAFGK  
 FLWEWASARFSWLSLLVPFVQWVGLSPTVWLSVIWMMWYGPSLYSILSPFLPLLPI  
 FFCLWVYI"

## ORIGIN

Query Match 95.4%; Score 1126.8; DB 10; Length 1170;  
 Best Local Similarity 97.7%; Pred. No. 0;  
 Matches 1143; Conservative 0; Mismatches 27; Indels 0; Gaps 0;

Qy	1	ATGGGGCAGAATCTTTCCACCAGCAATCCTCTGGGATTCTTTCCCGACCACCAGTTGGAT	60
Db	1	ATGGGACAGAATCTTTCCACCAGCAATCCTCTGGGATTCTTTCCCGACCACCAGTTGGAT	60
Qy	61	CCAGCCTTCAGAGCAAACACCAACAATCCAGATTGGGACTTCAATCCCAACAAGGACACC	120
Db	61	CCAGCCTTCAGAGCAAACACCGCAAATCCAGATTGGGACTTCAATCCCAACAAGGACACC	120
Qy	121	TGGCCAGACGCCAACAAGGTAGGAGCTGGAGCATTCGGACTGGGGTTCACCCACCGCAC	180
Db	121	TGGCCAGACGCCAACAAGGTAGGAGCTGGAGCATTCGGGCTGGGATTACCCACCGCAC	180
Qy	181	GGAGGCCTTTTGGGGTGGAGCCCTCAGGCTCAGGGCATAACACAAACCTTGCCAGCAAAT	240
Db	181	GGAGGCCTTTTGGGGTGGAGCCCTCAGGCTCAGGGCATCATACAAACTTTGCCAGCAAAT	240
Qy	241	CCGCCTCCTGCTTCCACCAATCGCCAGTCAGGAAGGCAGCCTACCCCGCTGTCTCCACCT	300
Db	241	CCGCCTCCTGCATCCACCAATCGCCAGTCAGGAAGGCAGCCTACCCCGCTGTCTCCACCT	300
Qy	301	TTGAGAAACACTCATCCTCAAGCCATGCAGTGGAACTCCACAACCTTCCACCAAACCTG	360
Db	301	TTGAGAAACACTCATCCTCAGGCCATGCAGTGGAACTCCACAACCTTCCACCAAACCTG	360
Qy	361	CAAGATCCCAGAGTGAGAGGTCTGTATTTCCCTGCTGGTGGCTCCAGTTCAGGAACAGTA	420
Db	361	CAAGATCCCAGGGTGAGAGGCCTGTATTTCCCTGCTGGTGGCTCCAGTTCAGGAACAGTA	420
Qy	421	AACCCTGTTCCGACTACTGTCTCTCCCATATCGTCAATCTTCTCGAGGATTGGGGACCCT	480
Db	421	AACCCTGTTCCGACTACTGTCTCTCCCATATCGTCAATCTTCTCGAGGATTGGGGACCCT	480
Qy	481	GCGCGGAACATGGAGAACATCACATCAGGATTCTAGGACCCCTGCTCGTGTACAGGCG	540
Db	481	GCGCTGAACATGGAGAACATCACATCAGGATTCTAGGACCCCTGCTCGTGTACAGGCG	540
Qy	541	GGGTTTTTCTTGTGACAAGAATCCTCACAATACCGCAGAGTCTAGACTCGTGGTGGACT	600
Db	541	GGGTTTTTCTTGTGACAAGAATCCTCACAATACCGCAGAGTCTAGACTCGTGGTGGACT	600
Qy	601	TCTCTCAATTTTCTAGGGGGAATAACCGTGTGTCTTGGCCAAAATTTCGCAGTCCCCAACC	660
Db	601	TCTCTCAATTTTCTAGGGGGAATAACCGTGTGTCTTGGCCAAAATTTCGCAGTCCCCAACC	660
Qy	661	TCCAATCACTACCAACCTCCTGTCCTCCAACCTTGTCTGGTTATCGCTGGATGTGTCTG	720
Db	661	TCCAATCACTACCAACCTCCTGTCCTCCAACCTTGTCTGGTTATCGCTGGATGTGTCTG	720
Qy	721	CGGCGTTTTATCATCTTCCTCTTCATCCTGCTGCTATGCCTCATCTTCTTGTGGTTCTT	780
Db	721	CGGCGTTTTATCATCTTCCTCTTCATCCTGCTGCTATGCCTCATCTTCTTGTGGTTCTT	780
Qy	781	CTGGACTATCAAGGTATGTTGCCCGTTGTCTCTAATTCCAGGATCTTCAACCACCAGC	840
Db	781	CTGGACTATCAAGGTATGTTGCCCGTCTGTCTCTAATTCCAGGATCTTCAACCACCAGC	840
Qy	841	ACGGGACCATGCAGAGCCTGCACGACTCCTGCTCAAGGAACCTCTATGTATCCCTCCTGT	900
Db	841	GCGGGACCATGCAGAACCTGCACGACTACTGCTCAAGGAACCTCTATGTATCCCTCCTGT	900
Qy	901	TGCTGTACAAAACCTTCGGATGGAACTGCACCTGTATTCCCATCCCATCATCCTGGGCT	960
Db	901	TGCTGTACAAAACCTTCGACGGAAATTGCACCTGTATTCCCATCCCATCATCCTGGGCT	960

```
Qy      961  TTCGGAATAATTCCTATGGGAGTGGGCCTCAGCCCGTTTCTCCTGGCTCAGTTTACTAGTG 1020
      |||
Db      961  TTCGGAATAATTCCTATGGGAGTGGGCCTCAGCCCGTTTCTCCTGGCTCAGTTTACTAGTG 1020

Qy     1021  CCATTTGTTTCAGTGGTTCGTAGGGCTTCCCCCACTGTTTGGCTTTCAGTTATATGGATG 1080
      ||
Db     1021  CCCTTTGTTTCAGTGGTTCGTAGGGCTTCCCCCACTGTTTGGCTTTCAGTTATATGGATG 1080

Qy     1081  ATGTTGTACTGGGGGCCAAGTCTGTACACCATCTTGAGTCCCTTTTTACCGCTGTTACCA 1140
      ||||
Db     1081  ATGTGGTATTGGGGGCCAAGTCTGTACAGCATCTTGAGTCCCTTTTTACCGCTGTTACCA 1140

Qy     1141  ATTTTCTTTTGTCTTTGGGTATACATTAA 1170
      |||||
Db     1141  ATTTTCTTCTGTCTTTGGGTATACATTAA 1170
```

## RESULT 10

DQ219811

LOCUS DQ219811 3207 bp DNA linear SYN 23-OCT-2005

DEFINITION Synthetic construct isolate L1.1 Hepatitis B virus precore  
(precore), polymerase-reverse transcriptase (Pol), PreS1 (PreS1),  
and X protein (X) genes, complete cds.

ACCESSION DQ219811

VERSION DQ219811.1 GI:77819763

KEYWORDS .

SOURCE synthetic construct

ORGANISM synthetic construct  
other sequences; artificial sequences.

REFERENCE 1 (bases 1 to 3207)

AUTHORS Thakur,V., Kazim,S.N., Guptan,R.C., Malhotra,V. and Sarin,S.K.

TITLE Molecular epidemiology and transmission of hepatitis B virus in  
close family contacts of HBV-related chronic liver disease patients

JOURNAL J. Med. Virol. 70 (4), 520-528 (2003)

PUBMED 12794713

REFERENCE 2 (bases 1 to 3207)

AUTHORS Chakraborty,A.K., Chauhan,R. and Sarin,S.K.

TITLE A consensus method for numbering and representing Hepatitis B Virus  
genome useful for recombinant vector design

JOURNAL Unpublished

REFERENCE 3 (bases 1 to 3207)

AUTHORS Chakraborty,A.K., Chauhan,R. and Sarin,S.K.

TITLE Direct Submission

JOURNAL Submitted (21-SEP-2005) Gastroenterology, G. B. Pant Hospital,  
ICMR-Advanced Center for Liver Diseases, New Delhi, Delhi 110002,  
India

FEATURES Location/Qualifiers

```
source      1. .3207
              /organism="synthetic construct"
              /mol_type="other DNA"
              /serotype="ayw"
              /isolate="L1.1"
              /db_xref="taxon:32630"
              /country="India"
              /note="derived from Hepatitis B virus isolated from
              patient with chronic HBV infection
              genotype: D"

gene        1. .639
              /gene="precore"

CDS         1. .639
              /gene="precore"
              /note="contains core protein and eAg"
              /codon_start=1
```

```

/transl_table=11
/product="precore"
/protein_id="ABB04014.1"
/db_xref="GI:77819764"
/translation="MQLFHLCLIISCSCPTVQASKLCLGWLWGMDIDPYKEFGATVEL
LSFLPSDFFPSVRDLDLTASALYREALESPHCSPHHTALRQAILCWGELMTLATWVG
VNLEDPASRD LVSYVNTNMGLKFRQLLWFHISCLTFGRET VIEYLVSGVWIRTPPA
YRPPNAPILSTLPETTVVRRRGRSPRRRTPSPRRRRSQSPRRRRSQSRESQC"
gene 494. .2992
      /gene="Pol"
CDS   494. .2992
      /gene="Pol"
      /note="multifunctional protein having both DNA polymerase
and reverse transcriptase activity"
      /codon_start=1
      /transl_table=11
      /product="polymerase-reverse transcriptase"
      /protein_id="ABB04015.1"
      /db_xref="GI:77819765"
      /translation="MPLSYQHFRLLLLDDEAGPLEEELPRLADEGLNRRVAEDNLG
NLNVSIPWTHKVGNTGLYSSTVPVFNPHWKTPSFPNIHLHQDI IKKCEQFVGPLTVN
EKRRLLQLIMPARFYPKVTKYLPDKGIKPYYPEHLVNHYFQTRHYLHTLWKAGILYKR
ETTHSASFCGSPYSWEQDLLHGAESFHQQSSGILSRPPVGSSLQSKHRKSRLGLQSQQ
GHLARRQQGRSWSIRAGFHP TARRPFGVEPSGSGHTTNFASKSASCLHQS PVRKAAYP
AVSTFEKHSSSGHAVELHNLPPNSARSQSERP VFPCWWLQFRNSKPCSDYCLSLIVNL
LEDWGPCAEGHEHHIRI PRTPSRVTGGVFLVDKNPHNTAESRLVVD FSQFSRGN YRS
WPKFAVPNLQSLTNLLSSNLSWLSVDVSAAFHHLPLHPAAMP HLLVGSSGLSRYVARL
SSNSRILNNQHGTMPDLHDYCSRNLYVSLLLLYQTFGRKLHLYSHPI ILGFRKI PMGV
GLSPFLLAQFTSAICSVVRAFP HCLAFSYMDDVVLGAKSVQHLES LFTAVTNFLLSL
GIHLNPNKTKRWGYS LNFMGYVIGCYGSLPQEHI IQKIKECFRKL PINRPIDWKVCQR
IVGLLGFAAPFTQCGYPALMPLYACIQSKQAF TFSPTYKAFLCKQYLNLYPVARQPG
LCQVFADATPTGWGLVMGHQMRGTFSAPLP IHTAELLAACFARSRSGANI IGTDNSV
VLSRKYTSFPWLLGCAANWILRGTSFVYVPSALNPADDP SRGRLGLSRPLLRLPFRPT
TGRTSLYADSPSPVSHLPDRVHFASPLHVAWRPP"
gene 1035. .2204
      /gene="PreS1"
CDS   1035. .2204
      /gene="PreS1"
      /note="surface protein; antigenic protein; virus entry and
cell surface recognition"
      /codon_start=1
      /transl_table=11
      /product="PreS1"
      /protein_id="ABB04016.1"
      /db_xref="GI:77819766"
      /translation="MGQNLSTSNPLGFFPDHQLDPAFRANTANPDWDFNPNKDTWPD
ANKVGAGAFGLGFTPPHGGLLGWSPQAQGILQTL PANPPPASTNRQSGRQPTPLSPPLR
NTHPQAMQWNSTTFHQTLQDPRVRGLYFPAGGSSSGTVNPVLT TASPLSSIFSRIGDP
ALNMENITSGFLGPLLVLQAGFLLTRILTIPQSLDSWWTSLNFLGGTTVCLGQNSQS
PTSNHSPTSCPPTCPGYRWMCLRRFII FLFILLCLIFLLVLLDYQGM LPVCPLIPGS
STTSTGPCRTCMTTAQGTSMYPSCCCTKPSDGNCTCIPI PSSWAFGKFLWEWASARFS
WLSLLVPFVQWFVGLSPTVWLSVIWMWYWGPSLYSILSPFLPLLPIFFCLWVYI"
gene 2743. .3207
      /gene="X"
CDS   2743. .3207
      /gene="X"
      /note="transactivator protein; role in viral pathogenesis
including hepatocellular carcinoma; transactivates many
cellular genes"
      /codon_start=1
      /transl_table=11
      /product="X protein"
      /protein_id="ABB04017.1"
      /db_xref="GI:77819767"

```



```

/translation="MAARLCCQLDPARDVLCRLPVGAESCGRPFSGSLGTLSSPSPSA
VPTDHGAHLSLRGLPVCAFSSAGPCALRFTSARRMETTVNAHQILPKVLHKRTLGLSA
MSTTDLEAYFKDCLFKDWHEELGEEIRLKVFLVGGCRHKLVCAPAPCNFF TSA"
misc_feature      3183. 3207
                  /gene="X"
                  /note="25 nt added to keep all reading frames in order;
                        useful for vector design and genome analysis"

```

## ORIGIN

```

Query Match      95.4%;  Score 1126.6;  DB 8;  Length 3207;
Best Local Similarity 97.1%;  Pred. No. 0;
Matches 1147;  Conservative 0;  Mismatches 34;  Indels 0;  Gaps 0;

```

```

Qy      1  ATGGGGCAGAATCTTTCCACCAGCAATCCTCTGGGATTCTTTCCCGACCACCAGTTGGAT 60
      |||
Db    1035 ATGGGGCAGAATCTTTCCACCAGCAATCCTCTGGGATTCTTTCCCGACCACCAGTTGGAT 1094

Qy      61  CCAGCCTTCAGAGCAAACACCAACAATCCAGATTGGGACTTCAATCCCAACAAGGACACC 120
      |||
Db    1095 CCAGCCTTCAGAGCAAACACCGCAAATCCAGATTGGGACTTCAATCCCAACAAGGACACC 1154

Qy     121  TGGCCAGACGCCAACAAGGTAGGAGCTGGAGCATTCGGACTGGGGTTACCCCACCGCAC 180
      |||
Db    1155 TGGCCAGACGCCAACAAGGTAGGAGCTGGAGCATTCGGGCTGGGGTTTACCCCACCGCAC 1214

Qy     181  GGAGGCCTTTTGGGGTGGAGCCCTCAGGCTCAGGGCATAACACAAACCTTGCCAGCAAAT 240
      |||
Db    1215 GGAGGCCTTTTGGGGTGGAGCCCTCAGGCTCAGGGCATACTACAAACTTTGCCAGCAAAT 1274

Qy     241  CCGCCTCCTGCTTCCACCAATCGCCAGTCAGGAAGGCAGCCTACCCCGCTGTCTCCACCT 300
      |||
Db    1275 CCGCCTCCTGCTTCCACCAATCGCCAGTCAGGAAGGCAGCCTACCCCGCTGTCTCCACCT 1334

Qy     301  TTGAGAAACACTCATCCTCAAGCCATGCAGTGGAAGTCCACAACCTTCCACCAAACCTCTG 360
      |||
Db    1335 TTGAGAAACACTCATCCTCAGGCCATGCAGTGGAAGTCCACAACCTTCCACCAAACCTCTG 1394

Qy     361  CAAGATCCCAGAGTGAGAGGTCTGTATTTCCCTGCTGGTGGCTCCAGTTCAGGAACAGTA 420
      |||
Db    1395 CAAGATCCCAGAGTGAGAGGCCTGTATTTCCCTGCTGGTGGCTCCAGTTCAGGAACAGTA 1454

Qy     421  AACCCTGTTCCGACTACTGTCTCTCCCATATCGTCAATCTTCTCGAGGATTGGGGACCCT 480
      |||
Db    1455 AACCCTGTTCTGACTACTGCCTCTCCCTTATCGTCAATCTTCTCGAGGATTGGGGACCCT 1514

Qy     481  GCGCGGAACATGGAGAACATCACATCAGGATTCTAGGACCCCTGCTCGTGTACAGGCG 540
      |||
Db    1515 GCGCTGAACATGGAGAACATCACATCAGGATTCTAGGACCCCTTCTCGTGTACAGGCG 1574

Qy     541  GGGTTTTTCTTGTTGACAAGAATCCTCACAATACCGCAGAGTCTAGACTCGTGGTGGACT 600
      |||
Db    1575 GGGTTTTTCTTGTTGACAAGAATCCTCACAATACCGCAGAGTCTAGACTCGTGGTGGACT 1634

Qy     601  TCTCTCAATTTTCTAGGGGGAACCTACCGTGTGTCTTGCCAAAATTTCGCAGTCCCCAACC 660
      |||
Db    1635 TCTCTCAATTTTCTAGGGGGAACCTACCGTGTGTCTTGCCAAAATTTCGCAGTCCCCAACC 1694

Qy     661  TCCAATCACTACCAACCTCCTGTCTCCAACCTGTCTGGTTATCGCTGGATGTGTCTG 720
      |||
Db    1695 TCCAATCACTACCAACCTCTTGTCTCCAACCTGTCTGGTTATCGGTGGATGTGTCTG 1754

Qy     721  CGGCGTTTTATCATCTTCTCTCATCCTGCTGCTATGCCTCATCTTCTTGTGGTTCTT 780
      |||
Db    1755 CGGCGTTTTATCATCTTCTCTCATCCTGCTGCTATGCCTCATCTTCTTGTGGTTCTT 1814

```

```

Qy      781 CTGGACTATCAAGGTATGTTGCCCGTTTGTCTCTAATTCCAGGATCTTCAACCACCAGC 840
          |||
Db      1815 CTGGACTATCAAGGTATGTTGCCCGTTTGTCTCTAATTCCAGGATCCTCAACAACCAGC 1874

Qy      841 ACGGGACCATGCAGAGCCTGCACGACTCCTGCTCAAGGAACCTCTATGTATCCCTCCTGT 900
          |||
Db      1875 ACGGGACCATGCCGGACCTGCATGACTACTGCTCAAGGAACCTCTATGTATCCCTCCTGT 1934

Qy      901 TGCTGTACAAAACCTTCGGATGGAACTGCACCTGTATTCCCATCCCATCATCCTGGGCT 960
          |||
Db      1935 TGCTGTACCAAACCTTCGGACGGAAATTGCACCTGTATTCCCATCCCATCATCCTGGGCT 1994

Qy      961 TTCGGAAAATTCCTATGGGAGTGGGCCTCAGCCCGTTTCTCCTGGCTCAGTTTACTAGTG 1020
          |||
Db      1995 TTCGGAAAATTCCTATGGGAGTGGGCCTCAGCCCGTTTCTCCTGGCTCAGTTTACTAGTG 2054

Qy      1021 CCATTTGTTTCAGTGGTTCGTAGGGCTTTCCCCCACTGTTTGGCTTTTCAGTTATATGGATG 1080
          |||
Db      2055 CCATTTGTTTCAGTGGTTCGTAGGGCTTTCCCCCACTGTTTGGCTTTTCAGTTATATGGATG 2114

Qy      1081 ATGTTGTACTGGGGGCCAAGTCTGTACACCATCTTGAGTCCCTTTTTACCGCTGTTACCA 1140
          |||
Db      2115 ATGTGGTATTGGGGGCCAAGTCTGTACAGCATCTTGAGTCCCTTTTTACCGCTGTTACCA 2174

Qy      1141 ATTTTCTTTTGTCTTTGGGTATACATTTAAACCCTAATAAA 1181
          |||
Db      2175 ATTTTCTTTTGTCTTTGGGTATACATTTAAACCCTAACAAA 2215

```

## RESULT 11

AB104722

LOCUS AB104722 1170 bp DNA linear VRL 26-JUN-2003

DEFINITION Hepatitis B virus s gene for pre-S and S protein, complete cds,  
isolate: EG33.

ACCESSION AB104722

VERSION AB104722.1 GI:32261204

KEYWORDS .

SOURCE Hepatitis B virus

ORGANISM Hepatitis B virus

Viruses; Retro-transcribing viruses; Hepadnaviridae;  
Orthohepadnavirus.

REFERENCE 1

AUTHORS Saudy,N., Sugauchi,F., Tanaka,Y., Suzuki,S., Aal,A.A., Zaid,M.A.,  
Agha,S. and Mizokami,M.TITLE Genotypes and phylogenetic characterization of hepatitis B and  
delta viruses in Egypt

JOURNAL J. Med. Virol. 70 (4), 529-536 (2003)

PUBMED 12794714

REFERENCE 2 (bases 1 to 1170)

AUTHORS Suzuki,S., Saudy,N., Sugauchi,F., Orito,E., Agha,S. and Mizokami,M.

TITLE Direct Submission

JOURNAL Submitted (26-FEB-2003) Seiji Suzuki, Nagoya City University  
Graduate School, Department of clinical Molecular Informative  
Medicine; Mizuho, Nagoya, Aichi 467-8601, Japan  
(E-mail:seijis@med.nagoya-cu.ac.jp, Tel:81-52-853-8292,  
Fax:81-52-842-0021)

FEATURES Location/Qualifiers

source 1. .1170

/organism="Hepatitis B virus"

/mol\_type="genomic DNA"

/isolate="EG33"

/db\_xref="taxon:10407"

/country="Egypt"

```

gene      /note="genotype D"
          1. .1170
          /gene="s"
CDS       1. .1170
          /gene="s"
          /codon_start=1
          /product="pre-S and S protein"
          /protein_id="BAC78517.1"
          /db_xref="GI:32261205"
          /translation="MGQNLSTSNPLGFFPDHQLDPAFRANTANPDWDFNPKNKDTWPDAN
          KVGAGAFGLGFTPPHGGLLGWSPQAQGIQLTLPTNPPPPASTNRQSGRQPTPLSPPLR
          NTHPQAMQWNSTTFHQTLQDPRVRGLYFPAGGSSSGTVNPVPTTVSHISSIFSRIGDP
          ALNMENITSGFLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGGTTVCLGQNSQS
          PTSNHSPTSCPPTCPGYRWMCLRRFIIIFLIFLLVLLDYQGMLPVCPLIPGS
          STTSTGPCRTCTTPAQGTSMPSCCCTKPSDGNCTCIPIPSSWAFGKFLWEWASARFS
          WLSLLVPFVQWFVGLSPTVWLSVIWMMWYWGPSLYSILSPFLPLLPIFFCLWVYI"

```

## ORIGIN

Query Match 95.3%; Score 1125.2; DB 10; Length 1170;  
 Best Local Similarity 97.6%; Pred. No. 0;  
 Matches 1142; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

```

Qy      1 ATGGGGCAGAATCTTTCCACCAGCAATCCTCTGGGATTCTTTCCCGACCACCAGTTGGAT 60
          |||
Db      1 ATGGGGCAGAATCTTTCCACCAGCAATCCTCTGGGATTCTTTCCCGACCACCAGTTGGAT 60

Qy     61 CCAGCCTTCAGAGCAAACACCAACAATCCAGATTGGGACTTCAATCCCAACAAGGACACC 120
          |||
Db     61 CCAGCCTTCAGAGCAAACACCGCAAATCCAGATTGGGACTTCAATCCCAACAAGGACACT 120

Qy    121 TGGCCAGACGCCAACAAGGTAGGAGCTGGAGCATTGCGACTGGGGTTACCCCCACCGCAC 180
          |||
Db    121 TGGCCAGACGCCAACAAGGTAGGAGCTGGAGCATTGCGGGCTGGGATTACCCCCACCGCAC 180

Qy    181 GGAGGCCTTTTGGGGTGGAGCCCTCAGGCTCAGGGCATAACACAAACCTTGCCAGCAAAT 240
          |||
Db    181 GGAGGTCTTTTGGGGTGGAGCCCTCAGGCTCAGGGCATACTACAAACCTTGCCAACAAAT 240

Qy    241 CCGCCTCCTGCTTCCACCAATCGCCAGTCAGGAAGGCAGCCTACCCCGCTGTCTCCACCT 300
          |||
Db    241 CCGCCTCCTGCCTCTACCAATCGCCAGTCAGGAAGGCAGCCTACCCCGCTGTCTCCACCT 300

Qy    301 TTGAGAAACACTCATCCTCAAGCCATGCAGTGGAAGTCCACAACCTTCCACCAAACCTG 360
          |||
Db    301 TTGAGAAACACACATCCTCAGGCCATGCAGTGGAAGTCCACTACCTTCCACCAAACCTG 360

Qy    361 CAAGATCCCAGAGTGAGAGGTCTGTATTTCCCTGCTGGTGGCTCCAGTTCAGGAACAGTA 420
          |||
Db    361 CAAGATCCCAGAGTGAGAGGCCTGTATTTCCCTGCTGGTGGCTCCAGTTCAGGAACAGTA 420

Qy    421 AACCTGTTCGACTACTGTCTCTCCCATATCGTCAATCTTCTCGAGGATTGGGGACCCT 480
          |||
Db    421 AACCTGTTCGACTACTGTCTCTCACATATCGTCAATCTTCTCGAGGATTGGGGACCCT 480

Qy    481 GCGCGGAACATGGAGAACATCACATCAGGATTCTAGGACCCCTGCTCGTGTACAGGCG 540
          |||
Db    481 GCGCTGAACATGGAGAACATCACATCAGGATTCTAGGACCCCTGCTCGTGTACAGGCG 540

Qy    541 GGGTTTTTCTTGTGACAGAATCCTCACAATACCGCAGAGTCTAGACTCGTGGTGGACT 600
          |||
Db    541 GGGTTTTTCTTGTGACAGAATCCTCACAATACCGCAGAGTCTAGACTCGTGGTGGACT 600

Qy    601 TCTCTCAATTTTCTAGGGGGAACACCGTGTGTCTTGGCCAAAATTTCGAGTCCCCAAC 660
          |||

```

```

Db      601 TCTCTCAATTTTCTAGGGGGAACCTACCGTGTGTCTTGGCCAAAATTCGCAGTCCCCAACC 660
Qy      661 TCCAATCACTCACCAACCTCCTGTCTCCAACCTTGTCTGGTTATCGCTGGATGTGTCTG 720
      |||
Db      661 TCCAATCACTCACCAACCTCCTGTCTCCAACCTTGTCTGGTTATCGCTGGATGTGTCTG 720
Qy      721 CGGCGTTTTATCATCTTCTCTTCATCCTGCTGCTATGCCTCATCTTCTTGTGGTTCTT 780
      |||
Db      721 CGGCGTTTTATCATCTTCTCTTCATCCTGCTGCTATGCCTCATCTTCTTGTGGTTCTT 780
Qy      781 CTGGACTATCAAGGTATGTTGCCCGTTTGTCTCTAATTCCAGGATCTTCAACCACCAGC 840
      |||
Db      781 CTGGACTATCAAGGTATGTTGCCCGTTTGTCTCTAATTCCAGGATCTTCAACTACCAGC 840
Qy      841 ACGGGACCATGCAGAGCCTGCACGACTCCTGCTCAAGGAACCTCTATGTATCCCTCCTGT 900
      |||
Db      841 ACGGGACCATGCAGAACCTGCACGACTCCTGCTCAAGGAACCTCTATGTATCCCTCCTGC 900
Qy      901 TGCTGTACAAAACCTTCGGATGGAACTGCACCTGTATTCCCATCCCATCATCCTGGGCT 960
      |||
Db      901 TGCTGTACAAAACCTTCGGACGGAAATGCACCTGTATTCCCATCCCATCATCCTGGGCT 960
Qy      961 TTCGGAAAATTCTATGGGAGTGGGCCTCAGCCCGTTTCTCCTGGCTCAGTTTACTAGTG 1020
      |||
Db      961 TTCGGAAAATTCTATGGGAGTGGGCCTCAGCCCGTTTCTCCTGGCTCAGTTTACTAGTG 1020
Qy      1021 CCATTTGTTTCAGTGGTTCGTAGGGCTTTCCCCCACTGTTTGGCTTTCAGTTATATGGATG 1080
      |||
Db      1021 CCATTTGTTTCAGTGGTTCGTAGGGCTTTCCCCCACTGTTTGGCTTTCAGTTATATGGATG 1080
Qy      1081 ATGTTGTACTGGGGGCAAGTCTGTACACCATCTTGAGTCCCTTTTTACCGCTGTTACCA 1140
      |||
Db      1081 ATGTGGTATTGGGGGCAAGTCTGTACAGCATCTTGAGTCCCTTTTTACCGCTGTTACCA 1140
Qy      1141 ATTTTCTTTTGTCTTTGGGTATACATTTAA 1170
      |||
Db      1141 ATTTTCTTTTGTCTTTGGGTATACATTTAA 1170

```

## RESULT 12

AY576427

LOCUS AY576427 1170 bp DNA linear VRL 01-APR-2005

DEFINITION Hepatitis B virus isolate 39 large S protein gene, partial cds.

ACCESSION AY576427

VERSION AY576427.1 GI:50953650

KEYWORDS .

SOURCE Hepatitis B virus

ORGANISM Hepatitis B virus

Viruses; Retro-transcribing viruses; Hepadnaviridae;  
Orthohepadnavirus.

REFERENCE 1 (bases 1 to 1170)

AUTHORS Kew,M.C., Kramvis,A., Yu,M.C. and Hodkinson,J.

TITLE Increased hepatocarcinogenic potential of hepatitis B virus  
genotype A in black Africans

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 1170)

AUTHORS Kramvis,A. and Kew,M.C.

TITLE Direct Submission

JOURNAL Submitted (19-MAR-2004) Medicine (Molecular Hepatology Research  
Unit), University of the Witwatersrand, 7 York Road, Johannesburg,  
GP 2193, South Africa

FEATURES Location/Qualifiers

source 1. .1170

/organism="Hepatitis B virus"

## CDS

```

/mol_type="genomic DNA"
/serotype="ayw2"
/isolate="39"
/db_xref="taxon:10407"
/country="South Africa"
/note="genotype: D"
<1. .1170
/note="envelope protein"
/codon_start=1
/product="large S protein"
/protein_id="AAT90407.1"
/db_xref="GI:50953651"
/translation="MGQNLSTSNPLGFFPDHQLDPAFRANTANPDWDFNPNKDTWPDAN
KVGAGAFGLGFTPPHGGLLGWSPQAQGIQLTPANPPPASTNRQSGRQPTPLSPPLR
NTHPQAMQWNSTTFHQTLDPRVRGLYFPAGSSSGTVNPVPTTASPLSSIFSRIGDP
ALNMENITSGFLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGGTTVCLGQNSQS
PTSNHSPTSCPPTCPGYRWMCLRRFIIFLFILLLCLIFLLVLLDYQGMLPVCPLIPGS
STTSTGPCRTCTTPAQGTSMPSCCCTKPSDGNCTCIPIPSSWAFGKFLWEWASARFS
WLSLLVPFVQWVGLSPTVWLSVIMMWYWGPSLYSILSPFLPLLPIFFCLWVYI"

```

## ORIGIN

Query Match 95.3%; Score 1125.2; DB 10; Length 1170;  
 Best Local Similarity 97.6%; Pred. No. 0;  
 Matches 1142; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

Qy	1	ATGGGGCAGAATCTTTCCACCAGCAATCCTCTGGGATTCTTTCCCGACCACCAGTTGGAT	60
Db	1	ATGGGGCAGAATCTTTCCACCAGCAATCCTCTGGGATTCTTTCCCGACCACCAGTTGGAT	60
Qy	61	CCAGCCTTCAGAGCAAACACCAACAATCCAGATTGGGACTTCAATCCCAACAAGGACACC	120
Db	61	CCAGCCTTCAGAGCAAACACCGCAAATCCAGATTGGGACTTCAATCCCAACAAGGACACC	120
Qy	121	TGGCCAGACGCCAACAAGGTAGGAGCTGGAGCATTCGGACTGGGGTTCACCCACCGCAC	180
Db	121	TGGCCAGACGCCAACAAGGTAGGAGCTGGAGCATTCGGGCTGGGTTTCACCCACCGCAC	180
Qy	181	GGAGGCCTTTTGGGGTGGAGCCCTCAGGCTCAGGGCATAACACAAACCTTGCCAGCAAAT	240
Db	181	GGAGGCCTTTTGGGGTGGAGCCCTCAGGCTCAGGGCATACTACAACTTTGCCAGCAAAT	240
Qy	241	CCGCCTCCTGCTTCCACCAATCGCCAGTCAGGAAGGCAGCCTACCCCGCTGTCTCCACCT	300
Db	241	CCGCCTCCTGCCTCCACCAATCGCCAGTCAGGAAGGCAGCCTACCCCGCTGTCTCCACCT	300
Qy	301	TTGAGAAACACTCATCCTCAAGCCATGCAGTGGAATCCACAACCTTCCACCAAACCTCTG	360
Db	301	TTGAGAAACACTCATCCTCAGGCCATGCAGTGGAATCCACAACCTTCCACCAAACCTCTG	360
Qy	361	CAAGATCCCAGAGTGAGAGGTCTGTATTTCCCTGCTGGTGGCTCCAGTTCAGGAACAGTA	420
Db	361	CAAGATCCCAGAGTGAGAGGCCTGTATTTCCCTGCTGGTGGCTCCAGTTCAGGAACAGTA	420
Qy	421	AACCCTGTTCCGACTACTGTCTCTCCCATATCGTCAATCTTCTCGAGGATTGGGGACCCT	480
Db	421	AACCCTGTTCCGACTACTGCCTTCCCTTATCGTCAATCTTCTCGAGGATTGGGGACCCT	480
Qy	481	GCGCGGAACATGGAGAACATCACATCAGGATTCCCTAGGACCCCTGCTCGTGTACAGGCG	540
Db	481	GCGCTGAACATGGAGAACATCACATCAGGATTCCCTAGGACCCCTTCTCGTGTACAGGCG	540
Qy	541	GGGTTTTTCTTGTGACAAGAATCCTCACAATACCGCAGAGTCTAGACTCGTGGTGGACT	600
Db	541	GGGTTTTTCTTGTGACAAGAATCCTCACAATACCGCAGAGTCTAGACTCGTGGTGGACT	600

```

Qy      601 TCTCTCAATTTTCTAGGGGGAACACCGTGTGTCTTGGCCAAAATTTCGCAGTCCCCAACC 660
        |||
Db      601 TCTCTCAATTTTCTAGGGGGAACACCGTGTGTCTTGGCCAAAATTTCGCAGTCCCCAACC 660

Qy      661 TCCAATCACTACCAACCTCCTGTCTCCAACCTTGTCTGGTTATCGCTGGATGTGTCTG 720
        |||
Db      661 TCCAATCACTACCAACCTCCTGTCTCCAACCTTGTCTGGTTATCGCTGGATGTGTCTG 720

Qy      721 CGGCGTTTTATCATCTTCCTCTTCATCCTGCTGCTATGCCTCATCTTCTTGTGGTTCTT 780
        |||
Db      721 CGGCGTTTTATCATCTTCCTCTTCATCCTGCTGCTATGCCTCATCTTCTTGTGGTTCTT 780

Qy      781 CTGGACTATCAAGGTATGTTGCCCGTTTGTCTCTAATTCCAGGATCTTCAACCACCAGC 840
        |||
Db      781 CTGGACTATCAAGGTATGTTGCCCGTGTGTCTCTAATTCCAGGATCTTCAACCACCAGC 840

Qy      841 ACGGGACCATGCAGAGCCTGCACGACTCCTGCTCAAGGAACCTCTATGTATCCCTCCTGT 900
        |||
Db      841 ACGGGACCATGCCGAACCTGCACGACTCCTGCTCAAGGAACCTCTATGTATCCCTCCTGT 900

Qy      901 TGCTGTACAAAACCTTCGGATGGAAACTGCACCTGTATTCCCATCCCATCATCCTGGGCT 960
        |||
Db      901 TGCTGTACAAAACCTTCGGACGGAATTGCACCTGTATTCCCATCCCATCATCCTGGGCT 960

Qy      961 TTCGGAAAATTCCTATGGGAGTGGGCCTCAGCCCGTTTCTCCTGGCTCAGTTTACTAGTG 1020
        |||
Db      961 TTCGGAAAATTCCTATGGGAGTGGGCCTCAGCCCGTTTCTCCTGGCTCAGTTTACTAGTG 1020

Qy      1021 CCATTTGTTTCAGTGGTTCGTAGGGCTTTCCCCCACTGTTTGGCTTTCAGTTATATGGATG 1080
        |||
Db      1021 CCATTTGTTTCAGTGGTTCGTAGGGCTTTCCCCCACTGTTTGGCTTTCAGTTATATGGATG 1080

Qy      1081 ATGTTGTACTGGGGGCCAAGTCTGTACACCATCTTGAGTCCCTTTTTACCGCTGTTACCA 1140
        |||
Db      1081 ATGTGGTATTGGGGGCCAAGTCTGTACAGCATCTTGAGTCCCTTTTTACCGCTGTTACCA 1140

Qy      1141 ATTTTCTTTTGTCTTTGGGTATACATTTAA 1170
        |||
Db      1141 ATTTTCTTTTGTCTTTGGGTATACATTTAA 1170

```

## RESULT 13

AY603460

LOCUS AY603460 1170 bp DNA linear VRL 19-APR-2005

DEFINITION Hepatitis B virus isolate 02T large S protein (S), middle S protein (S), and S protein (S) genes, complete cds.

ACCESSION AY603460

VERSION AY603460.1 GI:47499931

KEYWORDS .

SOURCE Hepatitis B virus

ORGANISM Hepatitis B virus

Viruses; Retro-transcribing viruses; Hepadnaviridae; Orthohepadnavirus.

REFERENCE 1 (bases 1 to 1170)

AUTHORS Sominskaya,I., Mihailova,M., Jansons,J., Emelyanova,V., Folkmane,I., Smagris,E., Dumpis,U., Rozentals,R. and Pumpens,P.

TITLE Hepatitis B and C virus variants in long-term immunosuppressed renal transplant patients in Latvia

JOURNAL Intervirology 48 (2-3), 192-200 (2005)

PUBMED 15812194

REFERENCE 2 (bases 1 to 1170)

AUTHORS Sominskaya,I., Mihailova,M., Jansons,J., Emelyanova,V., Folkmane,I., Smagris,E., Dumpis,U., Rozentals,R. and Pumpens,P.

TITLE Direct Submission  
JOURNAL Submitted (21-APR-2004) Protein Engineering Department, Biomedical Research and Study Centre, University of Latvia, Ratsupites str. 1, Riga LV-1067, Latvia

FEATURES Location/Qualifiers

source 1. .1170  
/organism="Hepatitis B virus"  
/virion  
/mol\_type="genomic DNA"  
/isolate="02T"  
/isolation\_source="renal transplantation patient"  
/db\_xref="taxon:10407"  
/country="Latvia"  
/note="subtype: ayw2; genotype: D"

gene 1. .1170  
/gene="S"

CDS 1. .1170  
/gene="S"  
/note="pre-S1/pre-S2/S; LHBS/MHBS/HBsAg; surface antigen"  
/codon\_start=1  
/product="large S protein"  
/protein\_id="AAT28717.1"  
/db\_xref="GI:47499932"  
/translation="MGQNLSTSNPLGFFPDHQLDPAFRANTANPDWDFNPNKDTWPDANKVGAGAFGLGFTPPHGGLLGWSPQAQGI IQTL PANPPPASTNRQSGRQPTPLSPPLRNTHPQAMQWNSTTFHQTLQDPRVRGLYFPAGGSSSGTVNPVPTTASPISSIFSRIGDPVLNMENITSGFLGPLLVLQAGFFLLTKIFTIPQSLDSWWTSLNFLGGTTVCLGQNSQSPSTNSHPTSCPPTCPGYRWMCLRRFII FLFILLLLCLIFLLVLLDYQGMLPVCPLIPGSSTTSTGPCRTCTTPAQGTSMYPPSCCCTKPSDGNCTCIPIPSSWAFGKFLWEWASARSSWLSLLVPFVQWFWGLSPTVWLSVIWMMWYWGPSLYSILSPFLPLLPIFFCLWVYI"

CDS 325. .1170  
/gene="S"  
/note="pre-S2/S; MHBS/HBsAg; surface antigen"  
/codon\_start=1  
/product="middle S protein"  
/protein\_id="AAT28718.1"  
/db\_xref="GI:47499933"  
/translation="MQWNSTTFHQTLQDPRVRGLYFPAGGSSSGTVNPVPTTASPISSIFSRIGDPVLNMENITSGFLGPLLVLQAGFFLLTKIFTIPQSLDSWWTSLNFLGGTTVCLGQNSQSPSTNSHPTSCPPTCPGYRWMCLRRFII FLFILLLLCLIFLLVLLDYQGMLPVCPLIPGSSTTSTGPCRTCTTPAQGTSMYPPSCCCTKPSDGNCTCIPIPSSWAFGKFLWEWASARSSWLSLLVPFVQWFWGLSPTVWLSVIWMMWYWGPSLYSILSPFLPLLPIFFCLWVYI"

CDS 490. .1170  
/gene="S"  
/note="HBsAg; surface antigen"  
/codon\_start=1  
/product="S protein"  
/protein\_id="AAT28719.1"  
/db\_xref="GI:47499934"  
/translation="MENITSGFLGPLLVLQAGFFLLTKIFTIPQSLDSWWTSLNFLGGTTVCLGQNSQSPSTNSHPTSCPPTCPGYRWMCLRRFII FLFILLLLCLIFLLVLLDYQGMPLPVCPLIPGSSTTSTGPCRTCTTPAQGTSMYPPSCCCTKPSDGNCTCIPIPSSWAFGKFLWEWASARSSWLSLLVPFVQWFWGLSPTVWLSVIWMMWYWGPSLYSILSPFLPLLPIFFCLWVYI"

## ORIGIN

Query Match 95.3%; Score 1125.2; DB 10; Length 1170;  
Best Local Similarity 97.6%; Pred. No. 0;  
Matches 1142; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

Qy 1 ATGGGGCAGAACTTTCCACCAGCAATCCTCTGGGATTCTTTCCCGACCACCAGTTGGAT 60  
|||||

Db 1 ATGGGGCAGAACTTTCCACCAGCAATCCTCTGGGATTCTTTCCCGACCACCAGTTGGAT 60

Qy	61	CCAGCCTTCAGAGCAAACACCAACAATCCAGATTGGGACTTCAATCCCAACAAGGACACC	120
Db	61	CCAGCCTTCAGAGCAAACACAGCAAATCCAGATTGGGACTTCAATCCCAACAAGGACACC	120
Qy	121	TGGCCAGACGCCAACAAGGTAGGAGCTGGAGCATTTCGGACTGGGGTTACCCCCACCGCAC	180
Db	121	TGGCCAGACGCCAACAAGGTAGGAGCTGGAGCATTTCGGGCTGGGATTACCCCCACCGCAC	180
Qy	181	GGAGGCCTTTTGGGGTGGAGCCCTCAGGCTCAGGGCATAACACAAACCTTGCCAGCAAAT	240
Db	181	GGAGGCCTTTTGGGGTGGAGCCCTCAGGCTCAGGGCATAATACAAACCTTGCCAGCAAAT	240
Qy	241	CCGCCTCCTGCTTCCACCAATCGCCAGTCAGGAAGGCAGCCTACCCCGCTGTCTCCACCT	300
Db	241	CCGCCTCCTGCATCTACCAATCGCCAGTCAGGAAGGCAGCCTACCCCGCTGTCTCCACCT	300
Qy	301	TTGAGAAACACTCATCCTCAAGCCATGCAGTGGAACCTCCACAACCTTCCACCAAACCTCTG	360
Db	301	TTGAGAAACACTCATCCTCAGGCCATGCAGTGGAACCTCCACAACCTTCCACCAAACCTCTG	360
Qy	361	CAAGATCCCAGAGTGAGAGGTCTGTATTTCCCTGCTGGTGGCTCCAGTTCAGGAACAGTA	420
Db	361	CAAGATCCCAGGGTGAGAGGCCTGTATTTCCCTGCTGGTGGCTCCAGTTCAGGAACAGTA	420
Qy	421	AACCCTGTTCCGACTACTGTCTCTCCCATATCGTCAATCTTCTCGAGGATTGGGGACCCT	480
Db	421	AACCCTGTTCCGACTACTGCCTCTCCCATATCGTCAATCTTCTCGAGGATTGGGGACCCT	480
Qy	481	GCGCGGAACATGGAGAACATCACATCAGGATTCTAGGACCCCTGCTCGTGTTACAGGCG	540
Db	481	GTGCTGAACATGGAGAACATCACATCAGGATTCTAGGACCCCTGCTCGTGTTACAGGCG	540
Qy	541	GGGTTTTTCTTGTTGACAAGAATCCTCACAATACCGCAGAGTCTAGACTCGTGGTGGACT	600
Db	541	GGGTTTTTCTTGTTGACAAAAATCTTACAATACCGCAGAGTCTAGACTCGTGGTGGACT	600
Qy	601	TCTCTCAATTTTCTAGGGGGAACACCGTGTGTCTTGGCCAAAATTTCGCAGTCCCCAACC	660
Db	601	TCTCTCAATTTTCTAGGGGGAACACCGTGTGTCTTGGCCAAAATTTCGCAGTCCCCAACC	660
Qy	661	TCCAATCACTACCAACCTCCTGTCTCCTCAACTTGTCTGGTTATCGCTGGATGTGTCTG	720
Db	661	TCCAATCACTACCAACCTCCTGTCTCCTCAACTTGTCTGGTTATCGCTGGATGTGTCTG	720
Qy	721	CGGCGTTTTATCATCTTCTCTTCATCCTGCTGCTATGCCTCATCTTCTTGTGGTTCTT	780
Db	721	CGGCGTTTTATCATCTTCTCTTCATCCTGCTGCTATGCCTCATCTTCTTGTGGTTCTT	780
Qy	781	CTGGACTATCAAGGTATGTTGCCCGTTTGTCTCTAATTCCAGGATCTTCAACCACCAGC	840
Db	781	CTGGACTATCAAGGTATGTTGCCCGTTTGTCTCTAATTCCAGGATCTTCAACTACCAGC	840
Qy	841	ACGGGACCATGCAGAGCCTGCACGACTCCTGCTCAAGGAACCTCTATGTATCCCTCCTGT	900
Db	841	ACGGGACCATGCAGAACCTGCACGACTCCTGCTCAAGGAACCTCTATGTATCCCTCCTGT	900
Qy	901	TGCTGTACAAAACCTTCGGATGGAACTGCACCTGTATTCCCATCCCATCATCTGGGCT	960
Db	901	TGCTGTACCAAACCTTCGGACGAAATTGCACCTGTATTCCCATCCCATCATCTGGGCT	960
Qy	961	TTCGGAAAATTCCTATGGGAGTGGGCCTCAGCCCGTTTCTCCTGGCTCAGTTTACTAGTG	1020
Db	961	TTCGGAAAATTCCTATGGGAGTGGGCCTCAGCCCGTTTCTCCTGGCTCAGTTTACTAGTG	1020



```

Qy      1021 CCATTTGTTTCAGTGGTTCGTAGGGCTTTCCCCCACTGTTTGGCTTTCAGTTATATGGATG 1080
          |||
Db      1021 CCATTTGTTTCAGTGGTTCGTAGGGCTTTCCCCCACTGTTTGGCTTTCAGTTATATGGATG 1080

Qy      1081 ATGTTGTACTGGGGGCAAGTCTGTACACCATCTTGAGTCCCTTTTACCGCTGTTACCA 1140
          |||
Db      1081 ATGTGGTATTGGGGGCAAGTCTGTACAGCATCTTGAGTCCCTTTTACCGCTGTTACCA 1140

Qy      1141 ATTTTCTTTTGTCTTTGGGTATACATTTAA 1170
          |||
Db      1141 ATTTTCTTCTGTCTTTGGGTATACATTTAA 1170

```

## RESULT 14

AY603464

LOCUS AY603464 1170 bp DNA linear VRL 19-APR-2005

DEFINITION Hepatitis B virus isolate 12T large S protein (S), middle S protein (S), and S protein (S) genes, complete cds.

ACCESSION AY603464

VERSION AY603464.1 GI:47499947

KEYWORDS .

SOURCE Hepatitis B virus

ORGANISM Hepatitis B virus  
Viruses; Retro-transcribing viruses; Hepadnaviridae;  
Orthohepadnavirus.

REFERENCE 1 (bases 1 to 1170)

AUTHORS Sominskaya,I., Mihailova,M., Jansons,J., Emelyanova,V.,  
Folkmane,I., Smagris,E., Dumpis,U., Rozentals,R. and Pumpens,P.TITLE Hepatitis B and C virus variants in long-term immunosuppressed  
renal transplant patients in Latvia

JOURNAL Intervirology 48 (2-3), 192-200 (2005)

PUBMED 15812194

REFERENCE 2 (bases 1 to 1170)

AUTHORS Sominskaya,I., Mihailova,M., Jansons,J., Emelyanova,V.,  
Folkmane,I., Smagris,E., Dumpis,U., Rozentals,R. and Pumpens,P.

TITLE Direct Submission

JOURNAL Submitted (21-APR-2004) Protein Engineering Department, Biomedical  
Research and Study Centre, University of Latvia, Ratsupites str. 1,  
Riga LV-1067, Latvia

FEATURES Location/Qualifiers

source

```

1. .1170
/organism="Hepatitis B virus"
/virion
/mol_type="genomic DNA"
/isolate="12T"
/isolation_source="renal transplantation patient"
/db_xref="taxon:10407"
/country="Latvia"
/note="subtype: ayw3; genotype: D"

```

gene

```

1. .1170
/gene="S"

```

CDS

```

1. .1170
/gene="S"
/note="pre-S1/pre-S2/S; LHBS/MHBS/HBsAg; surface antigen"
/codon_start=1
/product="large S protein"
/protein_id="AAT28729.1"
/db_xref="GI:47499948"
/translation="MGQNLSTSNPLGFFPDHQLDPAFRANTANPDWDFNPNKDTWPDAN
KVGAGAFGLGFTPPHGGLLGWSPQAQGI IQTL PANPPPASTNRQSGRQPTPLSPPLR
NTHPQAMQWNSTTFHQTLQDPRVRGLYFPAGGSSSGTVNPVPTTVSPISSIFSRIGDP
ALNMENITSGFLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGGTTVCLGQNSQS
PTSNHSPTSCPPTCPGYRWMCLRRFII FLFILLLLCLIFLLVLLDYQGMLPVCPLIPGS

```

CDS  
 STTSAGTCRTCTTTAAGTSMYPSCCCTKPSDGNCTCIPIPSSWAFGKFLWEWASARFS  
 WLSLLVPFVQWVGLSPTVWLSVIWMMWYGPSLYSILSPFLPLLPIFFCLWVYI"  
 325. .1170  
 /gene="S"  
 /note="pre-S2/S; MHBS/HBsAg; surface antigen"  
 /codon\_start=1  
 /product="middle S protein"  
 /protein\_id="AAT28730.1"  
 /db\_xref="GI:47499949"  
 /translation="MQWNSTTFHQTLQDPRVRGLYFPAGGSSSGTVNPVPTTVSPISS  
 IFSRIGDPALNMENITSGFLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGGTTP  
 CLGQNSQSPTSNSHPTSCPPTCPGYRWMCLRRFIIIFLIFLLVLLDYQGMPL  
 VCPLIPGSSTTSAGTCRTCTTTAAGTSMYPSCCCTKPSDGNCTCIPIPSSWAFGKFLW  
 EWASARFSWLSLLVPFVQWVGLSPTVWLSVIWMMWYGPSLYSILSPFLPLLPIFFC  
 LWVYI"  
 CDS  
 490. .1170  
 /gene="S"  
 /note="HBsAg; surface antigen"  
 /codon\_start=1  
 /product="S protein"  
 /protein\_id="AAT28731.1"  
 /db\_xref="GI:47499950"  
 /translation="MENITSGFLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGG  
 TTVCLGQNSQSPTSNSHPTSCPPTCPGYRWMCLRRFIIIFLIFLLVLLDYQG  
 MLPVCPLIPGSSTTSAGTCRTCTTTAAGTSMYPSCCCTKPSDGNCTCIPIPSSWAFGK  
 FLWEWASARFSWLSLLVPFVQWVGLSPTVWLSVIWMMWYGPSLYSILSPFLPLLPI  
 FFCLWVYI"

## ORIGIN

Query Match 95.3%; Score 1125.2; DB 10; Length 1170;  
 Best Local Similarity 97.6%; Pred. No. 0;  
 Matches 1142; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

Qy	1	ATGGGGCAGAATCTTTCCACCAGCAATCCTCTGGGATTCTTTCCCGACCACCAGTTGGAT	60
Db	1	ATGGGGCAGAATCTTTCCACCAGCAATCCTCTGGGATTCTTTCCCGACCACCAGTTGGAT	60
Qy	61	CCAGCCTTCAGAGCAAACACCAACAATCCAGATTGGGACTTCAATCCCAACAAGGACACC	120
Db	61	CCAGCCTTCAGAGCAAACACCGCAAATCCAGATTGGGACTTCAATCCCAACAAGGACACC	120
Qy	121	TGGCCAGACGCCAACAAGGTAGGAGCTGGAGCATTCGGACTGGGGTTACCCCACCGCAC	180
Db	121	TGGCCAGACGCCAACAAGGTAGGAGCTGGAGCATTCGGGCTGGGATTACCCCACCGCAC	180
Qy	181	GGAGGCCTTTTGGGGTGGAGCCCTCAGGCTCAGGGCATAACACAAACCTTGCCAGCAAAT	240
Db	181	GGAGGCCTTTTGGGGTGGAGCCCTCAGGCTCAGGGCATCATAAACTTTGCCAGCAAAT	240
Qy	241	CCGCCTCCTGCTTCCACCAATCGCCAGTCAGGAAGGCAGCCTACCCGCTGTCTCCACCT	300
Db	241	CCGCCTCCTGCATCCACCAATCGCCAGTCAGGAAGGCAGCCTACCCGCTGTCTCCACCT	300
Qy	301	TTGAGAAACACTCATCCTCAAGCCATGCAGTGGAAGTCCACAACCTTTCCACCAAACCTCTG	360
Db	301	TTGAGAAACACTCATCCTCAGGCCATGCAGTGGAAGTCCACAACCTTTCCACCAAACCTCTG	360
Qy	361	CAAGATCCCAGAGTGAGAGGTCTGTATTTCCCTGCTGGTGGCTCCAGTTCAGGAACAGTA	420
Db	361	CAAGATCCCAGGGTGAGAGGCCTGTATTTCCCTGCTGGTGGCTCCAGTTCAGGAACAGTA	420
Qy	421	AACCCTGTTCCGACTACTGTCTCTCCCATATCGTCAATCTTCTCGAGGATTGGGGACCCT	480
Db	421	AACCCTGTTCCGACTACTGTCTCTCCCATATCGTCAATCTTCTCGAGGATTGGGGACCCT	480

Qy	481	GCGCGGAACATGGAGAACATCACATCAGGATTCTTAGGACCCCTGCTCGTGTGTACAGGCG	540
Db	481	GCGCTGAACATGGAGAACATCACATCAGGATTCTTAGGACCCCTGCTCGTGTGTACAGGCG	540
Qy	541	GGGTTTTTCTTGTGACAAGAATCCTCACAAATACCGCAGAGTCTAGACTCGTGGTGGACT	600
Db	541	GGGTTTTTCTTGTGACAAGAATCCTCACAAATACCGCAGAGTCTAGACTCGTGGTGGACT	600
Qy	601	TCTCTCAATTTCTAGGGGGAACACCGTGTGTCTTGGCCAAAATTCGCAGTCCCCAACC	660
Db	601	TCTCTCAATTTCTAGGGGGAACACCGTGTGTCTTGGCCAAAATTCGCAGTCCCCAACC	660
Qy	661	TCCAATCACTCACCAACCTCCTGTCTCCAACCTGTCTGGTTATCGCTGGATGTGTCTG	720
Db	661	TCCAATCACTCACCAACCTCCTGTCTCCAACCTGTCTGGTTATCGCTGGATGTGTCTG	720
Qy	721	CGGCGTTTTATCATCTTCTCTTCATCCTGCTGCTATGCCTCATCTTCTTGTGGTTCTT	780
Db	721	CGGCGTTTTATCATCTTCTCTTCATCCTGCTGCTATGCCTCATCTTCTTGTGGTTCTT	780
Qy	781	CTGGACTATCAAGGTATGTTGCCCGTTTGTCTCTAATTCCAGGATCTTCAACCACCAGC	840
Db	781	CTGGACTATCAAGGTATGTTGCCCGTCTGTCTCTAATTCCAGGATCTTCAACCACCAGC	840
Qy	841	ACGGGACCATGCAGAGCCTGCACGACTCCTGCTCAAGGAACCTCTATGTATCCCTCCTGT	900
Db	841	GCGGGAACATGCAGAACCTGCACGACTACTGCTCAAGGAACCTCTATGTATCCCTCCTGT	900
Qy	901	TGCTGTACAAAACCTTCGGATGGAAACTGCACCTGTATTCCCATCCCATCATCCTGGGCT	960
Db	901	TGCTGTACAAAACCTTCGGACGGAAATTGCACCTGTATTCCCATCCCATCATCCTGGGCT	960
Qy	961	TTCGGAAAATTCCTATGGGAGTGGGCCTCAGCCCGTTTCTCCTGGCTCAGTTTACTAGTG	1020
Db	961	TTCGGAAAATTCCTATGGGAGTGGGCCTCAGCCCGTTTCTCCTGGCTCAGTTTACTAGTG	1020
Qy	1021	CCATTTGTTTCACTGGTTTCGTAGGGCTTTCCCCCACTGTTTGGCTTTTCAATTATATGGATG	1080
Db	1021	CCCTTTGTTTCACTGGTTTCGTAGGGCTTTCCCCCACTGTTTGGCTTTTCAATTATATGGATG	1080
Qy	1081	ATGTTGTACTGGGGGCCAAGTCTGTACACCATCTTGAGTCCCTTTTTACCGCTCTTACCA	1140
Db	1081	ATGTGGTATTGGGGGCCAAGTCTGTACAGCATCTTGAGTCCCTTTTTACCGCTCTTACCA	1140
Qy	1141	ATTTTCTTTTGTCTTTGGGTATACATTTAA	1170
Db	1141	ATTTTCTTCTGTCTTTGGGTATACATTTAA	1170

RESULT 15

AY230128

LOCUS	AY230128	1186 bp	DNA	linear	VRL 23-NOV-2003
-------	----------	---------	-----	--------	-----------------

DEFINITION Hepatitis B virus isolate case 24 non-tumor large surface protein, middle surface protein, and small surface protein genes, complete cds.

ACCESSION      AY230128

VERSION AY230128.1 GI:38374285

## KEYWORDS

SOURCE	Hepatitis B virus
--------	-------------------

ORGANISM Hepatitis B virus  
Viruses; Retro-transcribing viruses; Hepadnaviridae;  
Orthohepadnavirus.

REFERENCE 1 (bases 1 to 1186)

AUTHORS Raimondo,G., Pollicino,T. and Raffa,G.  
 TITLE Occult HBV in liver cancer  
 JOURNAL Unpublished  
 REFERENCE 2 (bases 1 to 1186)  
 AUTHORS Raimondo,G., Pollicino,T. and Raffa,G.  
 TITLE Direct Submission  
 JOURNAL Submitted (04-FEB-2003) Internal Medicine, University of Messina,  
 via Consolare Valeria, Messina 98124, Italy

FEATURES Location/Qualifiers  
 source 1. .1186  
 /organism="Hepatitis B virus"  
 /virion  
 /mol\_type="genomic DNA"  
 /isolate="case 24 non-tumor"  
 /db\_xref="taxon:10407"

CDS 1. .1170  
 /note="preS1; preS/S"  
 /codon\_start=1  
 /product="large surface protein"  
 /protein\_id="AAR19340.1"  
 /db\_xref="GI:38374286"  
 /translation="MGQNLSTSNPLGFFPDHQLDPAFRANTANPDWDFNPNKDTWPDA  
 NKVGAGAFGLGFTPPHGGLLGWSPQAQGILQTLPANPPPASTNRQSGRQPTPLSPPLR  
 NTHPQAMQWNSTTFHQTLQDPRVRGLYFPAGGSSSGTVNPNVLTASPLSSIFSRIGDP  
 ALNMENITSGFLGPLLVLQAGFFLLTRILTIPQCLDSWWTSLNFLGGTTVCLGQNSQS  
 PTSNHSPTSCPPTCPGYRWMCLRRFIIIFLIFLLVLLDYQGMLPVCPLIPGS  
 STTSTGPCRTCMTTAQGTSMYPSCCCTKPSDGNCTCIPIPSWAFGKFLWEWASARFS  
 WLSLLVPFVQWFVGLSPTVWLSVIWMMWYWGPSLYSILSPFLPLLPIFFCLWVYI"

CDS 325. .1170  
 /note="preS2/S"  
 /codon\_start=1  
 /product="middle surface protein"  
 /protein\_id="AAR19341.1"  
 /db\_xref="GI:38374287"  
 /translation="MQWNSTTFHQTLQDPRVRGLYFPAGGSSSGTVNPNVLTASPLSS  
 IFSRIGDPALNMENITSGFLGPLLVLQAGFFLLTRILTIPQCLDSWWTSLNFLGGTTV  
 CLGQNSQSPTSNHSPTSCPPTCPGYRWMCLRRFIIIFLIFLLVLLDYQGMLP  
 VCPLIPGSSTTSTGPCRTCMTTAQGTSMYPSCCCTKPSDGNCTCIPIPSWAFGKFLW  
 EWASARFSWLSLLVPFVQWFVGLSPTVWLSVIWMMWYWGPSLYSILSPFLPLLPIFFC  
 LWVYI"

CDS 490. .1170  
 /note="S"  
 /codon\_start=1  
 /product="small surface protein"  
 /protein\_id="AAR19342.1"  
 /db\_xref="GI:38374288"  
 /translation="MENITSGFLGPLLVLQAGFFLLTRILTIPQCLDSWWTSLNFLGG  
 TTVCLGQNSQSPTSNHSPTSCPPTCPGYRWMCLRRFIIIFLIFLLVLLDYQG  
 MLPVCPLIPGSSTTSTGPCRTCMTTAQGTSMYPSCCCTKPSDGNCTCIPIPSWAFGK  
 FLWEWASARFSWLSLLVPFVQWFVGLSPTVWLSVIWMMWYWGPSLYSILSPFLPLLPI  
 FFCLWVYI"

## ORIGIN

Query Match 95.3%; Score 1125; DB 10; Length 1186;  
 Best Local Similarity 97.0%; Pred. No. 0;  
 Matches 1146; Conservative 0; Mismatches 35; Indels 0; Gaps 0;

Qy 1 ATGGGGCAGAATCTTTCCACCAGCAATCCTCTGGGATTCTTTCCCGACCACCAGTTGGAT 60  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 1 ATGGGGCAGAATCTTTCCACCAGCAATCCTCTGGGATTCTTTCCCGACCACCAGTTGGAT 60

Qy 61 CCAGCCTTCAGAGCAAACACCAACAATCCAGATTGGGACTTCAATCCCAACAAGGACACC 120  
 ||||||||||||||||||||| ||||||||||||||||||||||||||||  
 Db 61 CCAGCCTTCAGAGCAAACACCGCAAATCCAGATTGGGACTTCAATCCCAACAAGGACACC 120

Qy	121	TGGCCAGACGCCAACAAAGGTAGGAGCTGGAGCATTTCGGACTGGGGTTCACCCACCGCAC	180
Db	121	TGGCCAGACGCCAACAAAGGTAGGAGCTGGAGCATTTCGGGCTGGGTTTCACCCACCGCAC	180
Qy	181	GGAGGCCTTTTGGGGTGGAGCCCTCAGGCTCAGGGCATAACACAAACCTTGCCAGCAAAT	240
Db	181	GGAGGCCTTTTGGGGTGGAGCCCTCAGGCTCAGGGCATACTACAAACTTTGCCAGCAAAT	240
Qy	241	CCGCCTCCTGCTTCCACCAATCGCCAGTCAGGAAGGCAGCCTACCCCGCTGTCTCCACCT	300
Db	241	CCGCCTCCTGCTTCCACCAATCGCCAGTCAGGAAGGCAGCCTACCCCGCTGTCTCCACCT	300
Qy	301	TTGAGAAACACTCATCCTCAAGCCATGCAGTGGAATCCACAACCTTCCACCAAACCTCTG	360
Db	301	TTGAGAAACACTCACCTCAGGCCATGCAGTGGAATTCACAACCTTCCACCAAACCTCTG	360
Qy	361	CAAGATCCCAGAGTGAGAGGTCTGTATTTCCCTGCTGGTGGCTCCAGTTCAGGAACAGTA	420
Db	361	CAAGATCCCAGAGTGAGAGGCCTGTATTTCCCTGCTGGTGGCTCCAGTTCAGGAACAGTA	420
Qy	421	AACCCTGTTCCGACTACTGTCTCTCCCATATCGTCAATCTTCTCGAGGATTGGGGACCCT	480
Db	421	AACCCTGTTCTGACTACTGCCTCTCCCTATCGTCAATCTTCTCGAGGATTGGGGACCCT	480
Qy	481	GCGCGGAACATGGAGAACATCACATCAGGATTCCTAGGACCCCTGCTCGTGTTACAGGCG	540
Db	481	GCGCTGAACATGGAGAACATCACATCAGGATTCCTAGGACCCCTTCTCGTGTTACAGGCG	540
Qy	541	GGGTTTTTCTTGTTGACAAGAATCCTCACAATACCGCAGAGTCTAGACTCGTGGTGGACT	600
Db	541	GGGTTTTTCTTGTTGACAAGAATCCTCACAATACCGCAGTGTCTAGACTCGTGGTGGACT	600
Qy	601	TCTCTCAATTTTCTAGGGGGAACACCGTGTGTCTTGGCCAAAATTTCGCAGTCCCCAACC	660
Db	601	TCTCTCAATTTTCTAGGGGGAACACCGTGTGTCTTGGCCAAAATTTCGCAGTCCCCAACC	660
Qy	661	TCCAATCACTACCAACCTCCTGTCTCCTCAACTTGTCTGGTTATCGCTGGATGTGTCTG	720
Db	661	TCCAATCACTACCAACCTCTTGTCTCCTCAACTTGTCTGGTTATCGCTGGATGTGTCTG	720
Qy	721	CGGCGTTTTATCATCTTCCTCTTCATCCTGCTGCTATGCCTCATCTTCTTGTGGTTCTT	780
Db	721	CGGCGTTTTATCATCTTCCTCTTCATCCTGCTGCTATGCCTCATCTTCTTGTGGTTCTT	780
Qy	781	CTGGACTATCAAGGTATGTTGCCCGTTTGTCTCTAATTCCAGGATCTTCAACCACCAGC	840
Db	781	CTGGACTATCAAGGTATGTTGCCCGTTTGTCTCTAATTCCAGGATCCTCAACAACCAGC	840
Qy	841	ACGGGACCATGCAGAGCCTGCACGACTCCTGCTCAAGGAACCTCTATGTATCCCTCCTGT	900
Db	841	ACGGGACCATGCCGGACCTGCATGACTACTGCTCAAGGAACCTCTATGTATCCCTCCTGT	900
Qy	901	TGCTGTACAAAACCTTCGGATGGAAACTGCACCTGTATTCCCATCCCATCATCCTGGGCT	960
Db	901	TGCTGTACCAAACCTTCGGACGGAATTGCACCTGTATTCCCATCCCATCATCCTGGGCT	960
Qy	961	TTCGGAAAATTCTATGGGAGTGGGCCTCAGCCCGTTTCTCCTGGCTCAGTTTACTAGTG	1020
Db	961	TTCGGAAAATTCTATGGGAGTGGGCCTCAGCCCGTTTCTCCTGGCTCAGTTTACTAGTG	1020
Qy	1021	CCATTTGTTTCAGTGGTTTCGTAGGGCTTTCCCCCACTGTTTGGCTTTCAGTTATATGGATG	1080
Db	1021	CCATTTGTTTCAGTGGTTTCGTAGGGCTTTCCCCCACTGTTTGGCTTTCAGTTATATGGATG	1080

```

Qy      1081 ATGTTGTACTGGGGGCCAAGTCTGTACACCATCTTGAGTCCCTTTTTACCGCTGTTACCA 1140
        |||| ||| ||||||||||||||||||| ||||||||||||||||||| |||||||
Db      1081 ATGTGGTATTGGGGGCCAAGTCTGTACAGCATCTTGAGTCCCTTTTTACCGCTGTTACCA 1140

Qy      1141 ATTTTCTTTTGTCTTTGGGTATACATTAAACCCTAATAAA 1181
        ||||||||||||||||||| |||||
Db      1141 ATTTTCTTTTGTCTTTGGGTATACATTAAACCCTAACAAA 1181
    
```

Search completed: July 22, 2006, 07:03:19  
 Job time : 10973 secs

SCORE 1.3 BuildDate: 12/06/2005
---------------------------------